Binocular combination of luminance profiles

Jian Ding

School of Optometry and the Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA, USA

Dennis M. Levi

School of Optometry and the Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA, USA

We develop and test a new two-dimensional model for binocular combination of the two eyes' luminance profiles. For first-order stimuli, the model assumes that one eye's luminance profile first goes through a luminance compressor, receives gain-control and gain-enhancement from the other eye, and then linearly combines the other eye's output profile. For second-order stimuli, rectification is added in the signal path of the model before the binocular combination site. Both the total contrast and luminance energies, weighted sums over both the space and spatial-frequency domains, were used in the interocular gain-control, while only the total contrast energy was used in the interocular gain-enhancement. To challenge the model, we performed a binocular brightness matching experiment over a large range of background and target luminances. The target stimulus was a dichoptic disc with a sharp edge that has an increment or decrement luminance from its background. The disk's interocular luminance ratio varied from trial to trial. To refine the model we tested three luminance compressors, five nested binocular combination models (including the Ding–Sperling and the DSKL models), and examined the presence or absence of total luminance energy in the model. We found that (1) installing a luminance compressor, either a logarithmic luminance function or luminance gain-control, (2) including both contrast and luminance energies, and (3) adding interocular gain-enhancement (the DSKL model) to a combined model significantly improved its performance. The combined model provides a systematic account of binocular luminance summation over a large range of luminance input levels. It gives a unified explanation of Fechner's paradox observed on a dark background, and a winner-take-all phenomenon observed on a light background. To further test the model, we conducted two additional experiments: luminance summation of discs with asymmetric contour information (Experiment 2), similar to Levelt (1965) and binocular combination of second-order contrast-modulated gratings (Experiment 3). We used the model obtained in Experiment 1 to predict the results of Experiments 2 and 3 and the results of our previous studies. Model simulations further refined the contrast space weight and contrast sensitivity functions that are installed in the model, and provide a reasonable account for rebalancing of imbalanced binocular vision by reducing the mean luminance in the dominant eye.

Introduction

The human brain devotes enormous resources toward providing a cyclopean view of the world by combining the separate inputs from the two eyes. The process of binocular combination has been studied in a wide variety of different tasks including luminance change detection (Anstis & Ho, 1998; Baker, Wallis, Georgeson, & Meese, 2012; Cogan, 1987; Cohn & Lasley, 1976), contrast detection (Anderson & Movshon, 1989; Campbell & Green, 1965; Legge, 1984), contrast discrimination (Baker, Meese, & Georgeson, 2007; Ding & Levi, 2016; Georgeson, Wallis, Meese, & Baker, 2016; Legge, 1981, 1984; Meese, Georgeson, & Baker, 2006), contrast matching (Baker et al., 2007; Ding, Klein, & Levi, 2013b; Huang, Zhou, Zhou, & Lu, 2010; Legge & Rubin, 1981), Vernier acuity (Banton & Levi, 1991), orientation discrimination (Bearse & Freeman, 1994), visual direction (Mansfield & Legge, 1996), phase perception (Ding et al., 2013b; Ding & Sperling, 2006, 2007; Huang et al., 2010; Zhou, Georgeson, & Hess, 2014), and orientation perception (Yehezkel, Ding, Sterkin, Polat, & Levi, 2016). However, precisely how the brain combines the two eyes' images is still unclear. Typically, a model was developed specifically for one binocular task, but seldom addressed other tasks (Blake & Wilson, 2011). For example, Ding and Sperling (2006) proposed a gain control model to explain their phase data while Meese et al. (2006) proposed a two-stage model to explain

Citation: Ding, J., & Levi, D. M. (2017). Binocular combination of luminance profiles. *Journal of Vision*, *17*(13):4, 1–32, doi: 10.1167/17.13.4.

doi: 10.1167/17.13.4 Received April 18, 2017; published November 02, 2017 ISSN 1534-7362

their contrast discrimination data. Most models were only tested in a zero-dimensional space—with two numbers as inputs and one number as output rather than using two-dimensional (2D) images as the model's input and output.

The Ding–Sperling model was originally developed to account for the binocular combination of the two eyes' luminance profiles (2D), but it was only tested based on 0D phase data in a narrowly banded spatial frequency channel. In principle, it could be modified to explain other binocular combination tasks. Indeed, some efforts have been made to modify the model to address multiple binocular tasks (Ding et al., 2013b; Ding & Levi, 2016; Hou, Huang, Liang, Zhou, & Lu, 2013; Huang et al., 2010). By adding interocular contrast enhancement and a sensory fusion mechanism to the Ding–Sperling model (the DSKL model), we were able to successfully predict both binocular phase and contrast combination using one set of model parameters over a large range of input contrasts and phases in both normal (Ding et al., 2013b), and amblyopic observers (Ding, Klein, & Levi, 2013a). By inserting a contrast discrimination mechanism, either a nonlinear contrast transducer (NCT) or multiplicative noise (MN) into the model, we showed that the Ding–Sperling and the DSKL models could also predict binocular contrast discrimination (Ding & Levi, 2016). The best combination involved inserting early MN before the DSKL model, which successfully predicts the binocular advantage in contrast discrimination at high contrast levels. By inserting a binocular disparity energy mechanism, both the Ding–Sperling and DSKL models can also explain both disparity thresholds of Dmin and Dmax (Ding & Levi, 2016).

Interocular enhancement was first exposed in a study on amblyopic binocular vision (Ding et al., 2013b) where the interocular suppression is very asymmetric and the suppression from the nondominant eye (NDE) to the dominant eye (DE) is almost absent, thus revealing the NDE-to-DE's interocular enhancement. Interestingly, a recent physiological study (Shooner et al., 2017) confirmed our findings. They found asymmetric dichoptic masking in visual cortex of amblyopic macaque monkeys, and in the most severe amblyopes, they found the gain enhancement from the NDE to the DE. However, in normal symmetric binocular vision, the interocular enhancement is typically concealed by stronger interocular suppression and cannot be observed directly. In a modeling study, we demonstrated that adding interocular enhancement to a gain-control model results in significant improvement in model fitting in normal binocular phase and contrast combination (Ding et al., 2013a; Ding & Levi, 2016). More recently, adding interocular enhancement to the Ding–Sperling model successfully predicted binocular orientation combination (Yehezkel et al., 2016). The interocular enhancement correctly accounts for the reduced interocular suppression when base contrast is increased (Yehezhel et al., 2016).

So far, both the Ding–Sperling and the DSKL models have only been tested in experiments when the mean luminance of the two eyes was equal and remained constant, and therefore, no parameters reflect luminance variance. However, in the real world, the luminance might differ in the two eyes while viewing a three-dimensional (3D) surface and it may vary from time to time. In a recent study (Ding & Levi, 2014), we found that the interocular gain-control was dependent on the background luminance; reducing one eye's luminance reduces its suppression of the other eye, which makes it possible to rebalance binocular vision in people with asymmetric binocular vision (e.g., those with amblyopia) by wearing a neutral density filter in front of the dominant eye to reduce its suppression of the nondominant eye. Clearly, to account for binocular vision when the luminance differs in the two eyes, the luminance-modified interocular interactions should be added to a binocular combination model.

A luminance compressor, which allows the visual system to adapt to a wide range of luminance levels, has long been studied both psychophysically (Geisler, 1981, 1983; Hayhoe, Benimoff, & Hood, 1987; Hayhoe, Levin, & Koshel, 1992) and physiologically (Mante, Frazor, Bonin, Geisler, & Carandini, 2005; Yeh, Lee, & Kremers, 1996). However, remarkably no luminance compressor has ever been included in a binocular luminance summation model, although it is an essential process in the visual system. Depending on the background luminance, binocular luminance summation behaves very differently. For example, in binocular summation of luminance increments on a dark background, the equal perceived luminance data fall close to the linear contour over most of the range of interocular luminance ratios, but fold back to lower luminance excursions close to each axis (Engel, 1970; Levelt, 1965). This is Fechner's paradox—the observation that the appearance of unequal luminance in the two eyes seems dimmer than the brighter luminance viewed monocularly. In contrast, in the binocular summation of brightness decrements on a light background, the equal perceived luminance contour follows a winner-take-all rule (Anstis & Ho, 1998; Baker et al., 2012)—the perceived luminance decrement is the larger luminance decrement in the two eyes (similar to binocular contrast combination). Many alternative models have been proposed to account for luminance matching results (e.g., Anderson & Movshon, 1989; de Weert & Levelt, 1974; Engel, 1969; Grossberg & Kelly, 1999; Lehky, 1983), but none can explain these two different nonlinear phenomena simultaneously.

Recently, Baker et al. (2012) proposed a descriptive model for binocular brightness combination with one stimulus-dependent parameter, the power parameter (gamma), which is not fixed but dependent on the background luminance. Varying the gamma value produces a family of equal-brightness contours of differing curvature, which provides a good description of both Fechner's paradox and the winner-take-all phenomenon when using different gamma values. In a preliminary study (Ding & Levi, 2015), we added a luminance compressor (luminance gain-control) before the Ding–Sperling model, resulting in a systematic prediction of binocular luminance summation over a large range of luminance inputs.

Levelt (1965) reported several experiments on binocular summation of luminance discs on a dark background with or without a monocular contour. He found that the two eyes probably had equal weight in the summation when viewing the luminance discs without a concentric circle. However, when only one eye was presented with a monocular contour (a concentric circle), the eye with the contour dominated the summation, having more weight than the other eye. He also found that a local contour had more weight than a global contour, i.e., the target luminance had more weight when it was closer to the contour. Both the Ding–Sperling and DSKL models can give a reasonable account of this phenomenon—the eye with a concentric circle has more contrast energy and therefore has more weight in binocular luminance summation. However, without details of a more general model, precise predictions are impossible. In the current study we first develop a full 2D model for binocular combination of the two eyes luminance profiles, and then re-run the experiments of Levelt (1965—our Experiments 1 and 2) to see if the model simulation can give precise predictions.

Recently, Zhou et al. (2014) reported the "linear binocular combination" of responses to contrast modulation (CM) of second-order stimuli when their carrier contrast (CC) was identical (i.e., CC ratio = 1) in the two eyes. However, this linear phenomenon might be just a subset of a nonlinear system under special input conditions—the two eyes inputs are identical in the total contrast energy (TCE). Actually, all five nested nonlinear models, including the Ding–Sperling and DSKL models, tested in Ding et al. (2013b) predict that when the two eyes have equal TCE, they have equal weights in the binocular combination, which correctly predicts Zhou's data. Zhou et al. (2014) used a monocular probe (zero contrast in the other eye, i.e., CC ratio = 0) in a matching task, providing another data point to demonstrate that the ocular weights might vary with relative contrast in the two eyes. They noted that variations in relative contrast alter the weights assigned to each eye, but variations in modulation depth do not. They proposed a contrast-weighted model, in which the weight or gain assigned to each eye depends on the contrast in that eye relative to the other eye. The model successfully explained the linear binocular summation of CM depth when the two eyes have identical carrier contrast (i.e., CC ratio = 1), and can also explain the results when CC ratio = 0. However, it is not clear whether the model works at a CC ratio other than 1 or 0. To test whether a model, originally developed for first-order stimuli, also works for second-order stimuli, experiments should be performed with their carrier contrast varied (more than two points) in the two eyes. In Experiment 3 we examined binocular combination of second order stimuli when their carrier contrast differed in the two eyes, and simulated the full 2D model developed in this study to predict the results.

Methods

Stimuli

Experiment 1 used luminance discs (3° in diameter) with sharp edges as stimuli (Figures 1A–C). The target luminance was either increased against a dark or light background (white discs) or decreased against a light background (black discs). The stimulus duration was 200 ms and the interstimulus duration was 500 ms. The observer's task was to match the luminance increment or decrement from the background of the test disc to that of the standard disc. Test discs were presented with different target luminance in the two eyes; the standard discs had identical target luminance in the two eyes. The background luminance was the same for each eye, and the same for test and standard discs. The spatial layout of discs and contours (Figure 1A–F) was also the same for test and standard stimuli.

Stimuli were presented on a 22-inch NEC MultiSync CRT monitor (NEC, Tokyo, Japan) with a 1920×1440 spatial pixel resolution and 75 Hz vertical refresh rate. The experiments were controlled by a Mac Mini running MATLAB (MathWorks, Inc.) with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). A special circuit (Li, Lu, Xu, Jin, & Zhou, 2003) was used to yield 14 bit grayscale levels. Gamma correction was applied and verified by measuring 10 luminance levels using a Minolta LS-110 photometer (Minolta, Tokyo, Japan). The luminance of the monitor with all pixels set to the minimum value was 0.2 cd/m^2; the luminance with all pixels set to the maximum value was 74.2 cd/m^2. Displays were viewed in a mirror stereoscope and positioned optically 68 cm from the observer.

Experiment 2 used luminance discs with asymmetric contour information in the two eyes (Levelt, 1965). The

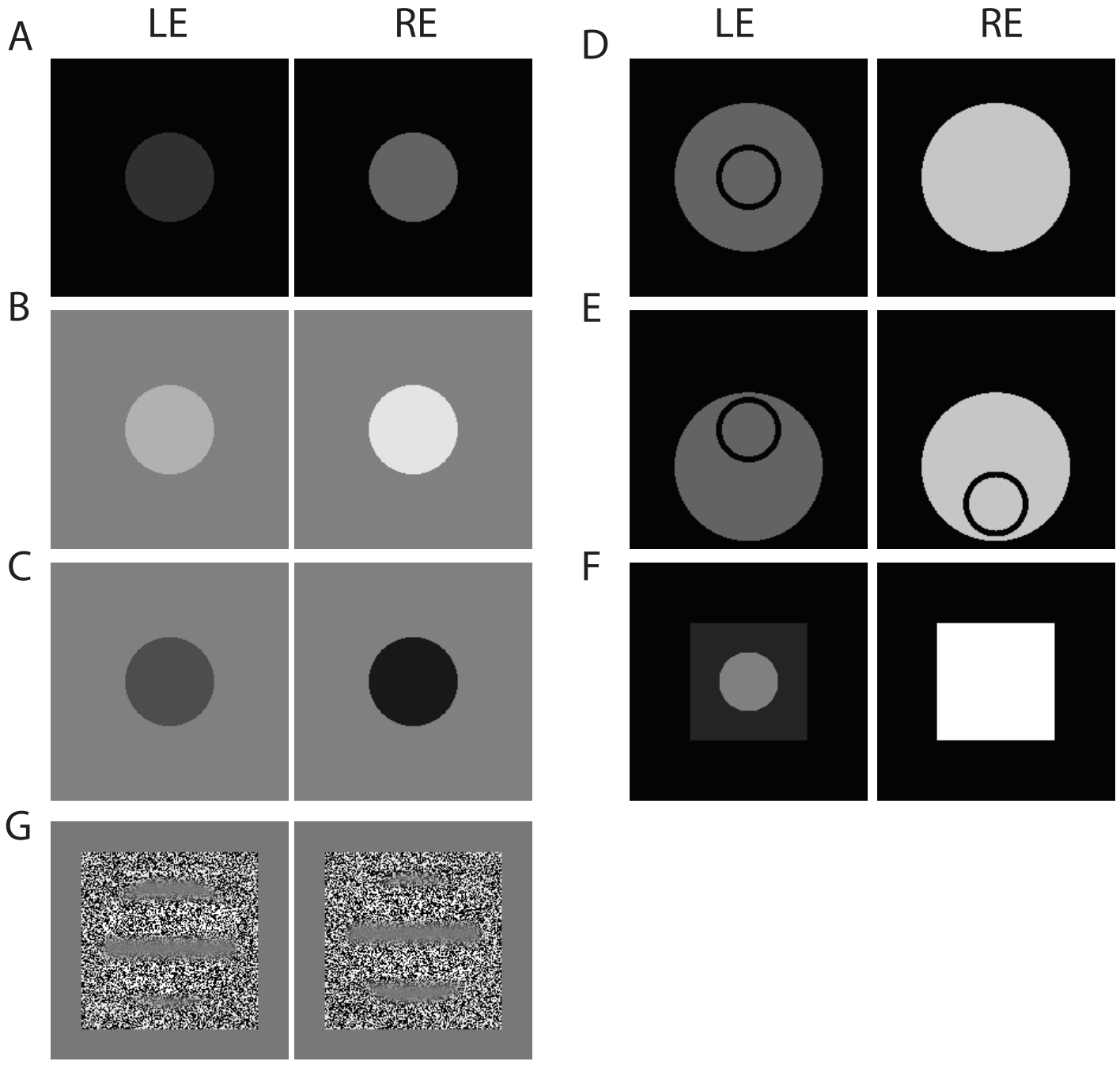

Figure 1. Sample stimuli presented to the left and right eyes (LE and RE). (A) A pair of discs with luminance increment from a dark background (Experiment 1a). (B) A pair of discs with luminance increment from a light background (Experiment 1a). (C) A pair of discs with luminance decrement from a light background (Experiment 1b). Test discs luminance differed in the two eyes, but the standard discs luminance was identical. (D) A pair of discs with luminance increment from a dark background with a concentric black circle only in the LE (Experiment 2a). (E) A pair of discs with luminance increment from a dark background with a black circle in each eye, one near the top of the LE disc and the other near the bottom of the RE disc (Experiment 2b). The target luminance was inside the top circle. (F) Stimuli for Experiment 2c. (G) Second-order contrast modulated (CM) gratings for Experiment 3 to measure the perceived CM phase and depth when interocular CM depth ratio (equal carrier contrast in the two eyes) or interocular carrier contrast ratio (equal CM depth in the two eyes) varied.

target luminance was increased from a dark background. In Experiment 2a (similar to experiment 2 in Levelt, 1965), the luminance discs were 5° in diameter with a black concentric circle (inner and outer diameters were 2° and 2.4°, respectively) either in the left or right eye. Figure 1D shows a pair of discs with a circle in the LE. In Experiment 2b (similar to experiment 3, in Levelt, 1965), the luminance disc and

circle were identical to those in Experiment 2a, except that each eye was presented with a black circle on a luminance disc, one near the top and the other the bottom (Figure 1E). The target luminance to be judged was the region inside the top circle. The other conditions were similar to Experiment 1. The stimulus duration was 200 ms and the interstimulus duration was 500 ms. The observer's task was to match the brightness of the test disc to that of the standard disc. Test discs were presented with different target luminance in the two eyes; the standard discs had identical target luminance in the two eyes.

In Experiment 2c (similar to experiment 4 in Levelt, 1965), one eye (e.g., RE) was presented only with a large luminance square (14° × 14°) and the other eye (e.g., LE) with a square (14° × 14°) plus a central disc that varied in size from trial to trial (Figure 1F). The target luminance was the central area of the disc. For a test, the luminance was fixed in the two eyes; the RE's square luminance was fixed at 72 cd/m^2, the LE's disc luminance was fixed at 12 cd/m^2 and its surrounding square luminance was fixed at 3.7 cd/m^2. For a standard, the RE's square and LE's central disc had identical luminance that was varied via a staircase to match a test luminance, and the LE's surrounding square luminance was always 1/3.25 of the luminance of the central disc. The background was always dark (0.2 cd/m^2). We tested seven different disc sizes (1°, 3°, 5°, 7°, 9°, 11°, 13° in diameters). The stimulus duration was 200 ms and the interstimulus duration was 500 ms.

Experiment 3 used second-order contrast modulated (CM) gratings as stimuli (Figure 1G) to measure the perceived CM phase (a) when the CM depth and phase differed in the two eyes but the carrier contrast (CC) was identical (=0.2), or (b) when the CC and the CM phase differed in the two eyes but the CM depth was identical (=0.7), and to measure the perceived CM depth (c) when the CM depth differed in the two eyes but the CM phase and CC (=0.2) were identical, or (d) when the CC differed in the two eyes but the CM depth (=0.7) and phase were identical. To make the results comparable with the previous study (Zhou et al., 2014), we also used a static, binary noise carrier for CM gratings. A horizontal CM sinewave grating was presented on a static, binary noise carrier (4.5° × 4.5° square with sharp edges), with identical spatial frequency (=0.68 cpd) of CM in the two eyes, circularly (3° in diameter) windowed with Gaussian-blurred edges, and duration of 1 s.

Procedure

For Experiment 1, there were two stimulus intervals: a standard, with equal luminance presented to the two eyes, and a test, with the interocular luminance ratio varying from trial to trial (RE/LE = 0, 0.125, 0.25, 0.5, 0.707, 1, 1.414, 2, 4, 8, or ∞). The test disc could be either in the first or second interval. For the luminance increment test there were five background luminances (0.2, 2.2, 4.2, 8.2, and 16.2 cd/m^2), and at each background luminance five luminance increments (1, 2, 4, 8, and16 cd/m^2) were tested. However, for the luminance decrement test, there were three background luminances (4.2, 8.2, and 16.2 cd/m^2). At 16.2 cd/m^2 background luminance, four luminance decrements (1, 2, 4, and 8 cd/m^2) were tested; at 8.2 cd/m^2, three luminance decrements (1, 2, and 4 cd/m^2) were tested; at 4.2 cd/m^2 , two luminance decrements (1 and 2 cd/m^2) were tested. For each pair of background luminance and luminance increment/decrement (luminance excursion) conditions, 22 one-up-one-down staircases (11 ratios × 2 presentation orders) were interleaved in one session to match the luminance increment or decrement from the background of a test disc to that of the standard disc. Each staircase contains 30 trials and a total of 660 trials were run for each session, which can be finished in about 1 hr. Observers were allowed to take short breaks inside the dark room during a session. The observer's task was to judge which interval contained a brighter disc for increment tests (white discs) or a darker disc for decrement tests (black discs). For each staircase, a psychometric function was fit to the response data, and the perceptually matched luminance was defined as the luminance that was equally likely to be judged above and below the standard luminance. The matched luminance was averaged across the presentation orders of test and standard discs.

The procedure for Experiment 2 was similar to Experiment 1. For Experiments 2a and 2b, 9 interocular luminance ratios (RE/LE = 0, 0.125, 0.25, 0.5, 1, 2, 4, 8, or ∞) at one luminance increment (8 cd/m^2) from a dark background (0.2 cd/m^2) were tested. In Experiment 2a, two sessions, each with 18 one-up-one-down staircases (9 ratios × 2 presentation orders), were run for the two conditions when a concentric circle was either in the left or right eye. Similarly, in Experiment 2b, two sessions were run for the two conditions when the top circle (target) was either in the left or right eye. For Experiment 2c, the test luminance was fixed in the two eyes while the disc size varied (see details in aforementioned Stimuli section). Two sessions, each with 14 one-up-one-down staircases (7 disc sizes × 2 presentation orders), were run for the two conditions when a luminance disc was either in the left or right eye. Each staircase consisted of 30 trials. The final results were averaged across presentation order and across the two eyes.

For Experiment 3 to measure the perceived CM phase and depth of a cyclopean CM grating, the

procedure was similar to our previous study with first order stimuli (Ding et al., 2013b). To measure the perceived phase, a horizontal reference line was attached to the sides of the stimuli. There was only one stimulus interval with one CM gratings present to each eye that either (a) differed in CM depth at a fixed equal CC or (b) differed in CC at a fixed equal CM depth. There were two phase orders in the two eyes: (1) LE phase = 45° and RE phase = −45°; (2) LE phase = −45°, and RE phase = 45°. Observers were asked to make a judgment of the position of the middle low contrast stripe relative to the attached reference line in a one-up-one-down staircase. One session with 18 one-up-one-down staircases (9 ratios × 2 phase orders) was run for each experiment. Each staircase consisted of 30 trials. The final results were averaged across phase orders. For the experiment to measure the perceived CM depth, there were two stimulus intervals, one with a reference CM grating only presented to one eye (the other eye is blank with mean luminance) and the other with test CM gratings presented to the two eyes, which either (c) differed in CM depth at a fixed equal CC or (d) differed in CC at a fixed equal CM depth. The observer was asked to match the perceived test CM depth to the reference by varying the reference in a one-up-one-down staircase. One session with 10 one-up-one-down staircases (5 ratios × 2 presentation orders) was run for each experiment. Each staircase contains 30 trials. The final results were averaged across the presentation orders.

Observers

Four observers with normal or corrected to normal vision participated in Experiment 1. The data were averaged across the four observers. Three observers with normal or corrected to normal vision participated in Experiment 2. The data were averaged across the three observers. Two observers with normal or corrected to normal vision participated in Experiment 3. The data were averaged across the two observers. All observers signed written consent forms.

Model

Let $I_L(x,y)$ and $I_R(x,y)$ be input luminance profiles in the two eyes. The binocular perceived profile is assumed to be the weighted summation of the two inputs, i.e.,

$$\hat{I}(x,y) = w_L(I_L, I_R)I_L(x,y) + w_R(I_L, I_R)I_R(x,y) \quad (1)$$

where the weights w_L and w_R depend on the two inputs. In principle, the perceived image features (e.g., brightness, contrast, contrast differences, contrast modulation depth, and phase) can be obtained from the perceived profile, making it possible to test the model across different tasks. Ding and Sperling (2006) proposed the first model with the property of Equation 1, in which the two weights were calculated according to gain-control theory. However, they only tested the model in binocular phase combination with a constant luminance background in a narrowly banded spatial frequency channel. Later, the Ding–Sperling model was further tested in other binocular tasks, such as phase and contrast, contrast discrimination, binocular disparity, brightness, and orientation, and was modified accordingly.

To predict binocular phase and contrast combination, we modified the Ding–Sperling model by adding interocular enhancement and a binocular fusion mechanism (the DSKL model; Ding et al., 2013b). We tested five nested models, including both the Ding–Sperling and the DSKL models, all with the property of Equation 1, and the DSKL model (the full model) gave the best fit for both phase and contrast data. However, because the tasks were performed on a constant luminance background, the models lack luminance components. In this study, we added luminance components to these five nested models to predict brightness perception of dichoptic luminance discs. Because the data in this study cannot be used to test the binocular fusion mechanism in the DSKL model, which was proposed to account for binocular contrast combination of two sinewave gratings with different phases, we did not include it in our current model.

Figure 2A shows the full model. One eye's luminance profile (e.g., $I_L(x,y)$) first goes through a luminance compressor (LOG: luminance logarithmic function or LG: luminance gain-control), and then after filtering though LoG (Laplacian of Gaussian) filters, the signal in each channel goes through a contrast-and-luminance gain-control-and-gain-enhancement unit (CLG), which receives gain-control and gain-enhancement from the other eye (say, the right eye). To avoid excessive interocular suppression and enhancement in binocular vision, the gain-control and gain-enhancement need to be constrained (Ding et al., 2013a, 2013b; Ding & Sperling, 2006, 2007), i.e., the RE's gain-control and gain-enhancement are gain-controlled by the left eye. Without gain-control of gain-control, the single-layer gain-control makes the binocular view much weaker than the monocular view under normal viewing conditions (high contrast), contradicting our daily experience (Ding & Sperling, 2006). Without gain-control of gain-enhancement, the single-layer gain-enhancement is very limited in improving model performance (Ding et al., 2013a, b; Ding & Levi, 2016). The product of the total weighted contrast energy and luminance energy (TCE × TLE) is used for gain-control and gain-control of gain-control. However, based on

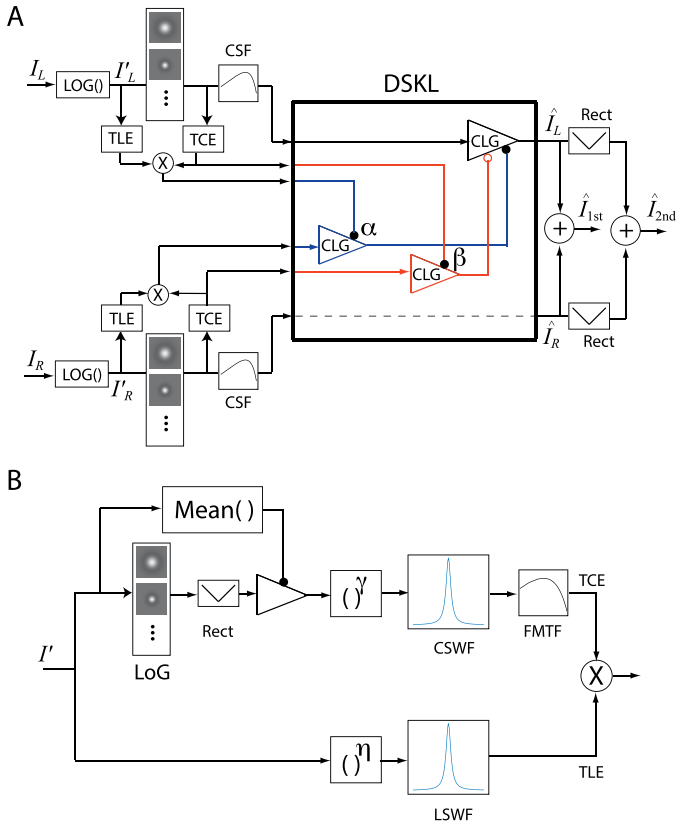

Figure 2. (A) Full model. A 2D luminance profile first goes through a luminance compressor (LOG: luminance logarithmic function), and then, after filtering through multiple LoG (Laplacian of Gaussian) filters, in each channel, one eye's (e.g., LE) signal receives contrast-and-luminance gain-control (CLG, blue path) and gain-enhancement (CLG, red path) from the other eye (e.g., RE). The gain-control depends on the product of the total contrast energy (TCE) and total luminance energy (TLE), which are both weighted sums over the space and spatial-frequency domains. The gain-enhancement only depends on the TCE. The RE's gain-control (blue) and gain-enhancement (red) themselves are gain-controlled by the LE with gain-control efficiency α and β respectively if we assume the gain-control efficiency $= 1$ in the LE's signal path (black). Only half of the DSKL model for LE's output is shown, and the other half for RE's output is symmetric to the LE's half. The first-order binocularly combined profile is the linear summation of the two monocular output profiles, and the second-order binocularly combined profile is the linear summation of those after rectification in each eye before binocular combination. (B) Calculation of TCE and TLE. For TCE (total contrast energy), a 2D compressive luminance profile is first filtered by LoG (Laplacian of Gaussian) filters, rectified, and then normalized by the mean luminance to give a contrast profile. For each spatial-frequency channel, the local contrast at each point in space is raised to the power of gamma, and then the contrast energy for this channel is calculated as a weighted sum over the two dimensional space (CSWF: contrast space weight function). The TCE is a weighted sum over spatial-frequency channels (FMTF: spatial-frequency →

modeling statistics (see Table 3 in the Modeling section), only the TCE is used for gain-enhancement and gain-control of gain-enhancement. Figure 2A shows half of the full model for the left eye, the other half for the right eye has a symmetric structure. For first-order perception, the binocular perceived profile is assumed to be the linear summation of two monocular output profiles, i.e., $\hat{I}(x,y) = \hat{I}_{\mathrm{L}}(x,y) + \hat{I}_{\mathrm{R}}(x,y)$. For second-order perception, rectification should be added in the signal path before binocular combination site to avoid the possible cancelation of outputs when two eyes carriers are anticorrelated, i.e., $\hat{I}(x,y) = \left|\hat{I}_{\mathrm{L}}(x,y)\right| + \left|\hat{I}_{\mathrm{R}}(x,y)\right|$. Please note that the rectification could be placed before the gain-control without changing the prediction of second-order binocular combination.

Different from our model that operates in 2D space along the whole signal path, the models proposed in Zhou et al. (2014) and Georgeson and Schofield (2016) assumed a FRF (filter-rectify-filter) structure for each eye to extract the second-order signal (a vector) from a 2D stimulus representation, and then, the two eyes' second-order signals were contrast-weighted and summed (vector summation) for binocular output.

Figure 2B shows how the total contrast energy (TCE) and total luminance energy (TLE) is calculated. For TCE, a 2D contrast profile is first obtained from the compressed luminance profile by filtering it with a LoG filter, rectifying and then normalizing with the mean luminance. For each spatial-frequency channel, the local contrast at each point in space is raised to the power of gamma, and then the contrast energy for this channel is calculated as a weighted sum over the two-dimensional space (contrast space weight function or CSWF). The TCE is a weighted sum of the contrast energy across all spatial-frequency channels (spatial-frequency modulation transfer function or FMTF). For TLE, the luminance value at each point in space of the compressed 2D luminance profile is raised to the power of eta, and then the TLE is calculated as a weighted sum over the two dimensional space (luminance space weight function or LSWF). Based on modeling statistics (see Table 3 in Modeling section), the TCE is used in both gain-control and gain-enhancement paths, whereas the TLE is only used in the gain-control path in the DSKL model.

←

modulation transfer function). For TLE (total luminance energy), the luminance value at each point in space of the compressed 2D luminance profile is raised to the power of eta, and then the TLE is calculated as a weighted sum over the two dimensional space (LSWF: luminance space weight function). The product of TCE and TLE is calculated for the gain-control path, whereas the gain-enhancement path only needs TCE.

Luminance compressor

To remain sensitive over a large range of input luminances, the visual system has developed a mechanism to compress the input luminance for brightness perception. We assume a local luminance compressor, applied to each pixel, which is supported by the experimental results of He and MacLeod (1998). They found that the light adaptation process had very high spatial resolution (>100 cpd), and suggested that it might be a strictly local luminance nonlinearity, one that either resides within individual photoreceptors or operates on signals from individual receptors. Here, we tested three luminance compressors, (1) the Stevens' power law for brightness perception, (2) a luminance gain-control, and (3) a luminance logarithmic function. Let $I(x, y)$ be a luminance profile and $I'(x, y)$ be the output of the compressor. The Stevens' power law is given by

$$I'(x,y) = I^p(x,y). \quad (2)$$

The luminance gain control is given by,

$$I'(x,y) = \frac{I^p(x,y)}{Z^q + I^q(x,y)}, \quad (3)$$

where Z is the gain-control threshold, and p and q are power parameters. When $I \gg Z$, the luminance gain control can be simplified to be the Stevens' power law for brightness perception. We also tested a luminance logarithmic function, given by,

$$I'(x,y) = \log\left(1 + \frac{I^p(x,y)}{Z^p}\right), \quad (4)$$

where Z is the threshold. When $I = 0$, we have $I' = 0$.

Total contrast energy (TCE) and total luminance energy (TLE)

Let I'_0 be the mean luminance of a profile after compressing, given by,

$$I'_0 = \frac{\int \hat{I}'(x,y) w_{\mathrm{Lum}}(x,y)\, dx\, dy}{\int w_{\mathrm{Lum}}(x,y)\, dx\, dy}. \quad (5)$$

where $w_{\mathrm{Lum}}(x, y)$ is a luminance space weighting function (LSWF) that should give more weight in the fovea for the calculation of mean luminance. The contrast at space point (x, y) is defined as its absolute Laplacian value normalized by the mean luminance, i.e.,

$$C(x,y) = \frac{1}{I'_0}|\Delta I'(x,y)| = \frac{1}{I'_0}\left|\frac{\partial^2 I'(x,y)}{\partial x^2} + \frac{\partial^2 I'(x,y)}{\partial y^2}\right|. \quad (6)$$

However, applying the Laplacian to a real image often produces high spatial frequency noise and needs to be smoothed by a Gaussian filter. Because both Gaussian and Laplacian are linear operators, people often exchange their order in the processing and use the LoG filter to estimate the Laplacian (Marr & Hildreth, 1980). For simplicity, it is common to use a DoG (Difference of Gaussian) filter to estimate the Laplacian. The visual system extracts image contrast in different scales (channels). A LoG filter in the ith channel is given by,

$$\mathrm{LoG}_i(x,y) = -\frac{1}{\pi\sigma_i^4}\left(1 - \frac{x^2+y^2}{2\sigma_i^2}\right)e^{-\frac{x^2+y^2}{2\sigma_i^2}}. \quad (7)$$

The contrast in the ith channel is given by the convolution of LoG_i and I', i.e.,

$$C_i(x,y) = \frac{1}{I'_0}(\mathrm{LoG}_i * I')(x,y)$$

The contrast energy in the ith channel is the weighted sum over the space, given by

$$\mathcal{E}_i(I) = b_i \int C_i^{\gamma}(x,y) w_{\mathrm{C}}(x,y)\, dx\, dy, \quad (8)$$

where $w_{\mathrm{C}}(x, y)$ is the contrast space weight functions (CSWF), b_i is a constant given by spatial frequency modulation transfer function (FMTF) (Ding & Sperling, 2006, 2007). The TCE is the weighted sum across all channels (Ding & Sperling, 2006, 2007), given by

$$\mathcal{E}(I) = \Sigma\, b_i \int C_i^{\gamma}(x,y) w_{\mathrm{C}}(x,y)\, dx\, dy. \quad (8')$$

The total luminance energy (TLE) is also the weighted sum over space, given by,

$$\mathcal{L}(I) = k \int I'^{\eta}(x,y) w_{\mathrm{Lum}}(x,y)\, dx\, dy, \quad (9)$$

where k is a constant. Figure 2B demonstrates how to calculate TCE and TLE.

In this study, we did not have enough data to test the details of the LoG filter, FMTF, and space weight functions, LSWF and CSWF. Instead, for the LoG filter, we estimated the contrast of an image with sharp edges only at one scale with $\sigma = 0.045°$, which reaches a peak at 10 cpd in the spatial frequency domain. With only one channel, the constant b_i given by FMTF can be selected to be a suitable value to simplify modeling (see the following). For the space weighting functions, we assumed that the weight reflects the density of photoreceptors in the retina, which reaches a peak in a small area (~0.032 deg^2; Curcio, Sloan, Kalina, & Hendrickson, 1990) in the fovea and falls steeply with increasing eccentricity (Curcio et al., 1990; Curcio, Sloan, Packer, Hendrickson, & Kalina, 1987; Perry & Cowey, 1985; Williams, 1988). It is given by

$$w_C(x,y) = w_{Lum}(x,y) = \frac{1}{1+\frac{1}{R_0^2}(x^2+y^2)}. \quad (10)$$

In the following, we show how to calculate the mean luminance, TCE and TLE, of a luminance disc. Let R_{disc} be radius of the disc, R_0 be the radius of a small area ($\sim 0.032^{\circ 2}$, i.e., $R_0 \approx 0.1^{\circ}$) where the weighting function reaches its peak, R_{max} be the radius of the maximum area that contributes for TCE and TLE, and I_T and I_B being the luminance of the disc (target) and its background respectively. From Equation 5, the mean luminance is given by,

$$I'_0 = \frac{\log\left(1+\left(\frac{R_{disc}}{R_0}\right)^2\right)}{\log\left(1+\left(\frac{R_{max}}{R_0}\right)^2\right)}\left(I'_T - I'_B\right) + I'_B.$$

For mathematical convenience, we assumed that,

$$\log\left(1+\left(\frac{R_{disc}}{R_0}\right)^2\right) = \frac{1}{2}\log\left(1+\left(\frac{R_{max}}{R_0}\right)^2\right), \quad (11)$$

which makes the target and background contribution to the mean luminance equal, i.e.,

$$I'_0 = \frac{I'_T + I'_B}{2}. \quad (12)$$

Similarly, for Condition (11), selecting a suitable value of k, the TLE in Equation 9 is given by,

$$\mathcal{L}(I) = \frac{I'^{\eta}_T + I'^{\eta}_B}{2}. \quad (13)$$

Also, for Condition (11), selecting a suitable value of b_i, the TCE in Equation 8 is given by,

$$\mathcal{E}(I) = \left(\frac{|I'_T - I'_B|}{I'_T + I'_B}\right)^{\gamma} = m^{\gamma}, \quad (14)$$

where m is the edge contrast of a disc, given by

$$m = \frac{|I'_T - I'_B|}{I'_T + I'_B}. \quad (15)$$

Please note that the quantity m given by Equation 15 is not simply the edge contrast in the screen image, but the one following luminance compression, and is therefore model-dependent to some extent.

Because the data were collected using discs with fixed size, R_0 and R_{max} were difficult to determine. Instead, we selected their values ($R_0 = 0.1^{\circ}$ and $R_{max} = 22.6^{\circ}$) to make Equation 11 valid for mathematical simplicity, so we can use Equations 13 and 14 to calculate TCE and TLE in the modeling. Without the assumption of Condition (11), the target and background luminances would have different contributions to the mean luminance, and TCE and TLE would require more complicated formulae, making model fitting more difficult. However, model simulation (not shown) shows that the model behaves in a similar fashion without the assumption of Condition (11), although some specific details might differ. The selection of a LoG filter was somewhat arbitrary. To extract the local contrast of an image with a sharp edge, we used a LoG filter with a relatively high peak spatial frequency (=10 cpd). For mathematical simplicity, we only extracted the local contrast at one scale. Model simulation (see the section of Model simulation) shows that extracting local contrast at a different scale has little effect on the model behavior, and the model behaves similarly when extracting local contrast at one or more scales (not shown).

We also assumed a flat contrast sensitivity function (CSF) for simplicity, so we did not have to split signals into multiple channels, process them separately, and then combine them for perception, but just combined the two eyes input profiles in the manner of Equation1 (see Equations 16–23 as follows). Installing a non-flat CSF has little effect on the model behavior when stimuli are simple (e.g., luminance discs with fixed size), but it much improves the model performance for complex stimuli (e.g., luminance discs with asymmetric contour information, see the Model simulation section).

Binocular combination models

In a previous study (Ding et al., 2013b), we proposed five nested models to predict both binocular phase and contrast combination. However, these models were developed for experiments when the mean luminance remained constant, and no model parameters reflected a luminance effect. However, the eye with reduced mean luminance (e.g., placing a neutral density filter before it) exerts less gain-control to the other eye (Ding & Levi, 2014). In the following, we modified these five nested models by including a luminance component.

Model 1: Contrast-and-luminance-weighted summation model (simplified Ding–Sperling model)

The Ding–Sperling model can be simplified to be a contrast-and-luminance-weighted summation model when the gain-control threshold = 0 or TCExTLE $\gg$ threshold. After binocular combination, the binocular output profile is given by,

$$\hat{I}(x,y) \approx \frac{\mathcal{E}_L(I_L)\mathcal{L}_L(I_L)}{\mathcal{E}_L(I_L)\mathcal{L}_L(I_L)+\mathcal{E}_R(I_R)\mathcal{L}_R(I_R)} I'_L(x,y) + \frac{\mathcal{E}_R(I_R)\mathcal{L}_R(I_R)}{\mathcal{E}_L(I_L)\mathcal{L}_L(I_L)+\mathcal{E}_R(I_R)\mathcal{L}_R(I_R)} I'_R(x,y). \quad (16)$$

where $I'_L(x,y)$ and $I'_R(x,y)$ are input luminance profiles after luminance compressing.

For a luminance disc, as noted already, at suitable values of R_0 ($\approx 0.1°$) and R_{max} ($\approx 22.6°$), after being normalized with the gain-control threshold g_c, its image contrast energy is given by,

$$\mathcal{E}_L = \left(\frac{m_L}{g_c}\right)^{\gamma} \quad \text{and} \quad \mathcal{E}_R = \left(\frac{m_R}{g_c}\right)^{\gamma}, \quad (17)$$

where m_L and m_R are the disc edge contrast in the two eyes given by Equation 15. Similarly, the image luminance energy of a luminance disc is given by

$$\mathcal{L}_L = \frac{{I'_{LT}}^{\eta} + {I'_{LB}}^{\eta}}{2} \quad \text{and} \quad \mathcal{L}_R = \frac{{I'_{RT}}^{\eta} + {I'_{RB}}^{\eta}}{2}. \quad (18)$$

Model 2: Ding–Sperling model with symmetric double-layer interocular gain-controls

The gain-control and the gain-control of gain-control have the same gain-control efficiency. After binocular combination, the binocular output profile is given by,

$$\hat{I}(x,y) = \frac{1}{1+\frac{\mathcal{E}_R(I_R)\mathcal{L}_R(I_R)}{1+\mathcal{E}_L(I_L)\mathcal{L}_L(I_L)}} I'_L(x,y) + \frac{1}{1+\frac{\mathcal{E}_L(I_L)\mathcal{L}_L(I_L)}{1+\mathcal{E}_R(I_R)\mathcal{L}_R(I_R)}} I'_R(x,y), \quad (19)$$

Model 3: Ding–Sperling model with asymmetric double-layer interocular gain-controls

The gain-control and the gain-control of gain-control have different gain-control efficiency. After binocular combination, the binocular output profile is given by,

$$\hat{I}(x,y) = \frac{1}{1+\frac{\mathcal{E}_R(I_R)\mathcal{L}_R(I_R)}{1+\alpha^{\gamma}\mathcal{E}_L(I_L)\mathcal{L}_L(I_L)}} I'_L(x,y) + \frac{1}{1+\frac{\mathcal{E}_L(I_L)\mathcal{L}_L(I_L)}{1+\alpha^{\gamma}\mathcal{E}_R(I_R)\mathcal{L}_R(I_R))}} I'_R(x,y), \quad (20)$$

where α is the gain-control efficiency, which is also raised to the power of γ.

Model 4: Adding interocular gain-enhancement to Model 3

After binocular combination, the binocular output profile is given by,

$$\hat{I}(x,y) = \frac{1+\mathcal{E}^*_R(I_R)}{1+\frac{\mathcal{E}_R(I_R)\mathcal{L}_R(I_R)}{1+\alpha^{\gamma}\mathcal{E}_L(I_L)\mathcal{L}_L(I_L)}} I'_L(x,y) + \frac{1+\mathcal{E}^*_L(I_L)}{1+\frac{\mathcal{E}_L(I_L)\mathcal{L}_L(I_L)}{1+\alpha^{\gamma}\mathcal{E}_R(I_R)\mathcal{L}_R(I_R)}} I'_R(x,y). \quad (21)$$

The image contrast energy for gain enhancement is given by

$$\mathcal{E}^*_L = \left(\frac{m_L}{g_e}\right)^{\gamma} \quad \text{and} \quad \mathcal{E}^*_R = \left(\frac{m_R}{g_e}\right)^{\gamma}, \quad (22)$$

where g_e is the contrast gain-enhancement threshold.

Model 5 (the full model): Adding interocular gain-control of gain-enhancement to Model 4

After binocular combination, the binocular output profile is given by,

$$\hat{I}(x,y) = \frac{1+\frac{\mathcal{E}^*_R(I_R)}{1+\beta^{\gamma}\mathcal{E}_L(I_L)}}{1+\frac{\mathcal{E}_R(I_R)\mathcal{L}_R(I_R)}{1+\alpha^{\gamma}\mathcal{E}_L(I_L)\mathcal{L}_L(I_L)}} I'_L(x,y) + \frac{1+\frac{\mathcal{E}^*_L(I_L)}{1+\beta^{\gamma}\mathcal{E}_R(I_R)}}{1+\frac{\mathcal{E}_L(I_L)\mathcal{L}_L(I_L)}{1+\alpha^{\gamma}\mathcal{E}_R(I_R)\mathcal{L}_R(I_R)}} I'_R(x,y), \quad (23)$$

where β is the gain-control efficiency, which is also raised to the power of γ. We note that the gain-enhancement and the gain-control of gain-enhancement only depend on the contrast energy. In the Modeling section, we test different model configurations with or without the luminance energy in gain-control and or gain-enhancement.

Results

Experiment 1: Binocular combination of luminance discs

For each session, the background luminance was identical in the two eyes and remained unchanged during the entire session, whereas the luminance of a test disc, increased or decreased from the background, differed in the two eyes, and varied from trial to trial. The background luminance was varied between sessions.

Experiment 1a: Equal-luminance-increment contour

At each interocular luminance ratio of a test disc, the luminance increment from the background was perceptually matched in brightness to that of the standard disc that was presented identically to the two eyes. Figure 3 shows equal-luminance-increment contours averaged across the four observers; each panel represents the contour for one pair of standard luminance increment (indicated at the top) and background luminance (indicated on the right). In the first row (panels at 0.2 cd/m^2 background luminance—the lowest luminance level of the display), when the standard luminance increment increases from 1 (the leftmost panel) to 16 (the most right panel) cd/m^2, Fechner's paradox becomes increasingly evident (i.e., the appearance of unequal luminance in the two eyes seems dimmer than the brighter luminance viewed monocularly). However, when the background luminance increases (from top to bottom), Fechner's paradox becomes less evident, and the winner-take-all phenomenon becomes more apparent when the background luminance was 16.2 cd/m^2 (bottom row). Fechner's paradox depends on both background and target luminance: It is most evident at the largest luminance increment from the black background (top-right panel). Over all, the shape of contours changes systematically depending on both background luminance and target luminance increment. More interestingly, the contour shape is very similar across all the panels within any given left-oblique diagonal, and changes progressively as you move in the right-oblique direction. This surely means that the main driver of contour shape (for increments) is the standard disc contrast (ratio of standard increment to background level) — very high at top right, and low at bottom left (thanks to Mark Georgeson for this observation and comment). After adding a luminance compressor and including luminance energy, the DSKL model was able to describe all these contours with different shapes as indicated by the fitting curves.

Experiment 1b: Equal-luminance-decrement contour

When the test luminance decreased from the background (e.g., a dark disc), its darkness was matched to the standard one. Figure 4 shows equal-luminance-decrement contours averaged across the four observers; each panel represents the contour for one pair of standard luminance decrement (indicated at the top) and background luminance (indicated on the right). Consistent with previous studies (Anstis & Ho, 1998; Baker et al., 2012), Fechner's paradox was not observed in equal-luminance-decrement contours. Compared with increment contours (Figure 3), the winner-take-all behavior is more evident in decrement contours, e.g., at 16.2 cd/m^2 background luminance, the winner-take-all behavior is more evident in the 8 cd/m^2 luminance decrement contour than the 8 cd/m^2 increment contour, which was correctly predicted by the model.

Experiment 2: Binocular combination of luminance discs with asymmetric contour information

To test the model with more complex stimuli, following Levelt (1965), we added asymmetric contours to the two eyes' luminance discs for binocular luminance summation. Because the formula of TCE and TLE were more complicated for a disc with extra contour information, it would be difficult, if not impossible, to fit the model directly to the data. Instead, we did model simulations using parameters already obtained in Experiment 1 to see how well the model can predict the data. Here, we tested three LoG filters (peak spatial frequency at 5, 10, and 20 cpd), three CSWFs (LSWF = CSWF was assumed), and two CSFs (flat and nonflat) in the simulation. Except for the one used in Experiment 1 (Equation 10), the other two CSWFs have different dropping rates in space weight when going away from the fovea, one slower and one more rapid than Equation 10, given by,

$$w_C(x,y) = w_{Lum}(x,y) = \frac{1}{1+\frac{1}{R_0}\sqrt{x^2+y^2}} \quad (24)$$

and

$$w_C(x,y) = w_{Lum}(x,y) = \frac{1}{1+\frac{1}{R_0^4}(x^2+y^2)^2}, \quad (25)$$

respectively.

In the simulation of Experiment 1 (binocular combination of two luminance discs without concentric circles; see examples in Figure 5A), we found that the predictions with different LoG filters or CSFs are almost identical (The dashed magenta, green, and black curves are overlapped with the thick black curve; and the dashed red curve is overlapped with the solid red curve in Figure 5A), and the predictions with different CSWFs (thick black curve versus solid blue and red curves) are also very similar.

Experiment 2a: The eye with more monocular contours dominates binocular combination

The model predicts that the eye with larger TCE dominates the binocular combination. To test this

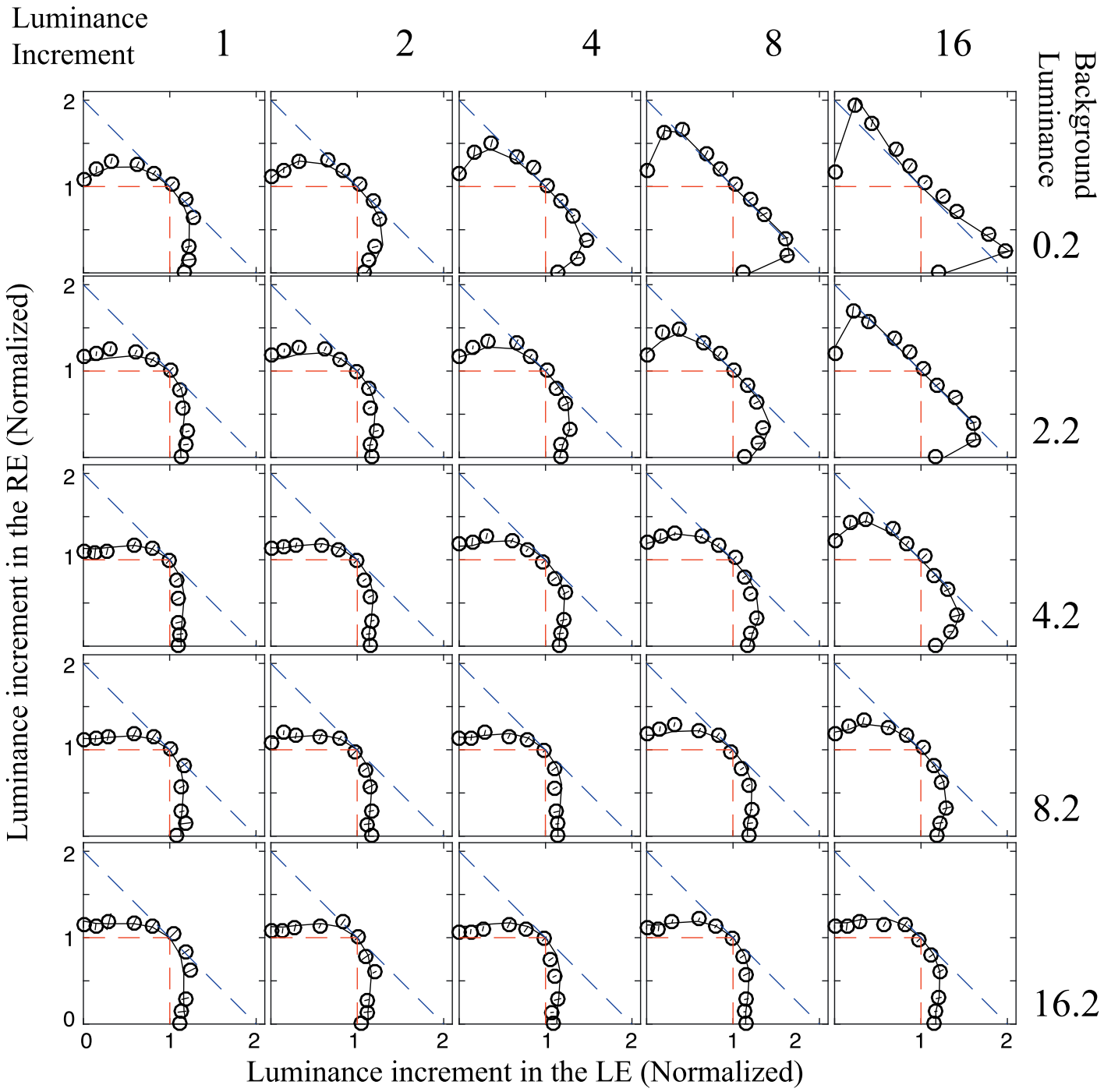

Figure 3. Equal-luminance-increment contour. The test luminance increment (normalized by the standard luminance increment) from the background (cd/m^2, as indicated in the right numbers) was matched to that of the standard luminance increment (cd/m^2, as indicated in the top numbers). For example, for the panel in the first column and second row, the background luminance is 2.2 cd/m^2 and the standard luminance increment is 1 cd/m^2 (e.g., the disc standard luminance is 3.2 cd/m^2). The black curves are fits of the full model shown in Figure 2. The red dashed lines show the predictions of the winner-take-all model, and the blue dashed lines shows the predictions of the linear summation model.

prediction, we added a black concentric circle to only one eye when performing the binocular luminance summation task. Consistent with Levelt (1965), we found that the eye with a circle dominated the binocular combination (Figure 5B), which confirmed the model predictions (smooth curves in Figure 5B) because the eye with a circle has larger TCE. Unlike the simulation of Experiment 1, adding a concentric circle to only one eye separated model predictions with different CSWFs or different CSFs, while the predic-

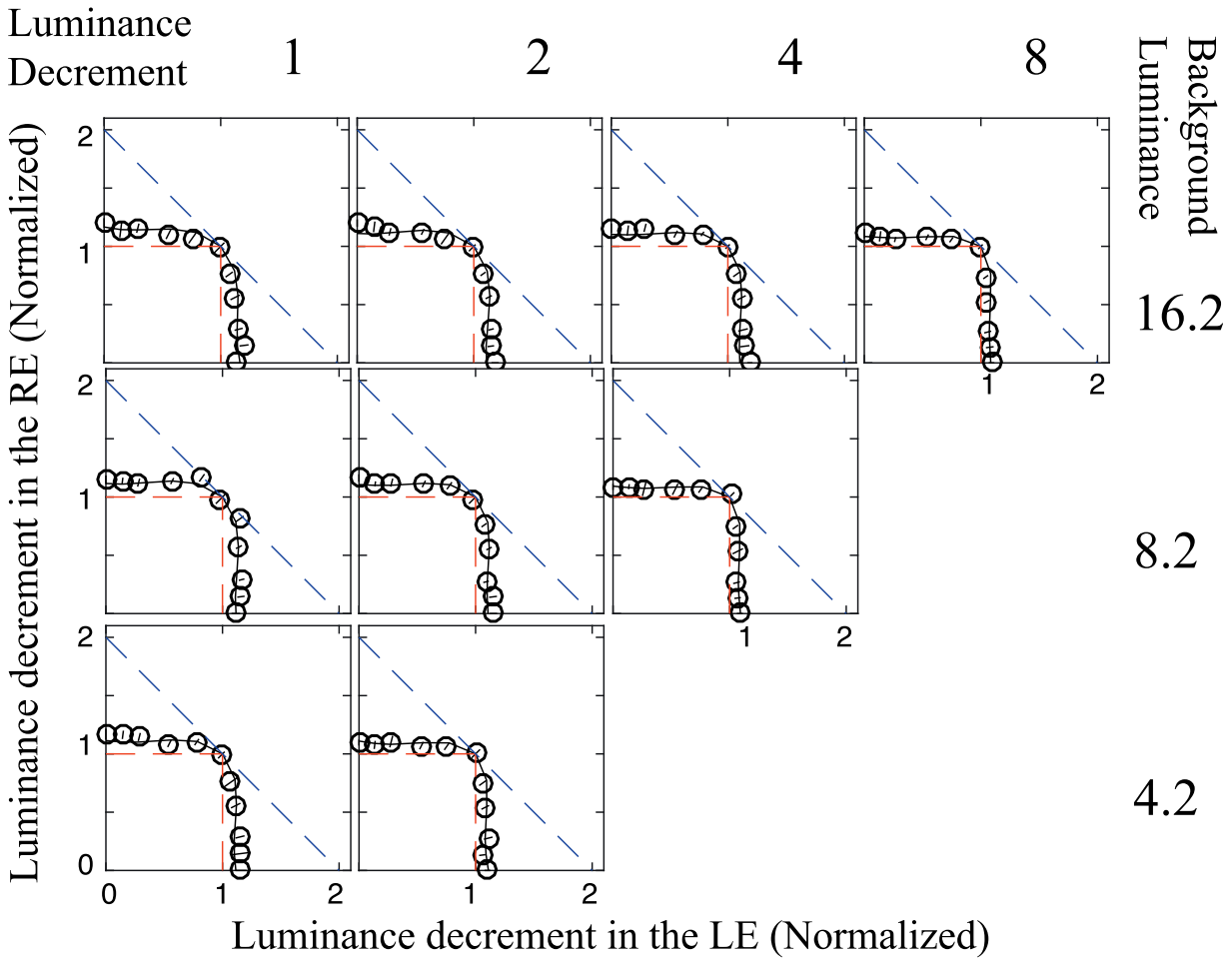

Figure 4. Equal-luminance-decrement contours. The test luminance decrement from the background (cd/m^2, as indicated in the right numbers) was matched to that of the standard luminance decrement (cd/m^2, as indicated in the top numbers). For example, for the panel in the first column and second row, the background luminance is 8.2 cd/m^2 and the standard luminance decrement is 1 cd/m^2 (e.g., the disc standard luminance is 7.2 cd/m^2). The black curves are fits of the full model shown in Figure 2. The red dashed lines show the predictions of the winner-take-all model, and the blue dashed lines shows the predictions of the linear summation model.

tions with different LoG filters remain similar. Installing a nonflat CSF provides better model predictions (dashed black and red curves) than flat CSF.

Experiment 2b: The eye weight in binocular combination depends on distance from the contour

The model also predicts that the eye with a near contour has a larger weight in the combination than with a distant contour because the space weight (CSWF) drops off sharply when away from the target. To address this prediction, we added two circles in the two eyes at two different locations of a luminance disc, one in the top half, and the other in the bottom half of the disc. The target was always the luminance inside the top circle. As predicted from the model (smooth curves in Figure 5C), the data (circle markers in Figure 5C) shows that the eye with top circle dominated the combination, consistent with Levelt (1965). Again, installing a nonflat CSF provides better model predictions (dashed black and red curves) than flat CSF, whereas models with different LoG filters give similar predictions. The CSWF given by Equation 25 (the space weight drops most sharply when away from the target) combined with a nonflat CSF has the best performance in model prediction (dashed red curve).

Experiment 2c: The dominant eye in binocular combination depends on monocular contour size

When the disc size decreases, its edge moves toward the fovea and therefore the contribution from its local edge to the TCE should increase (because of Equation

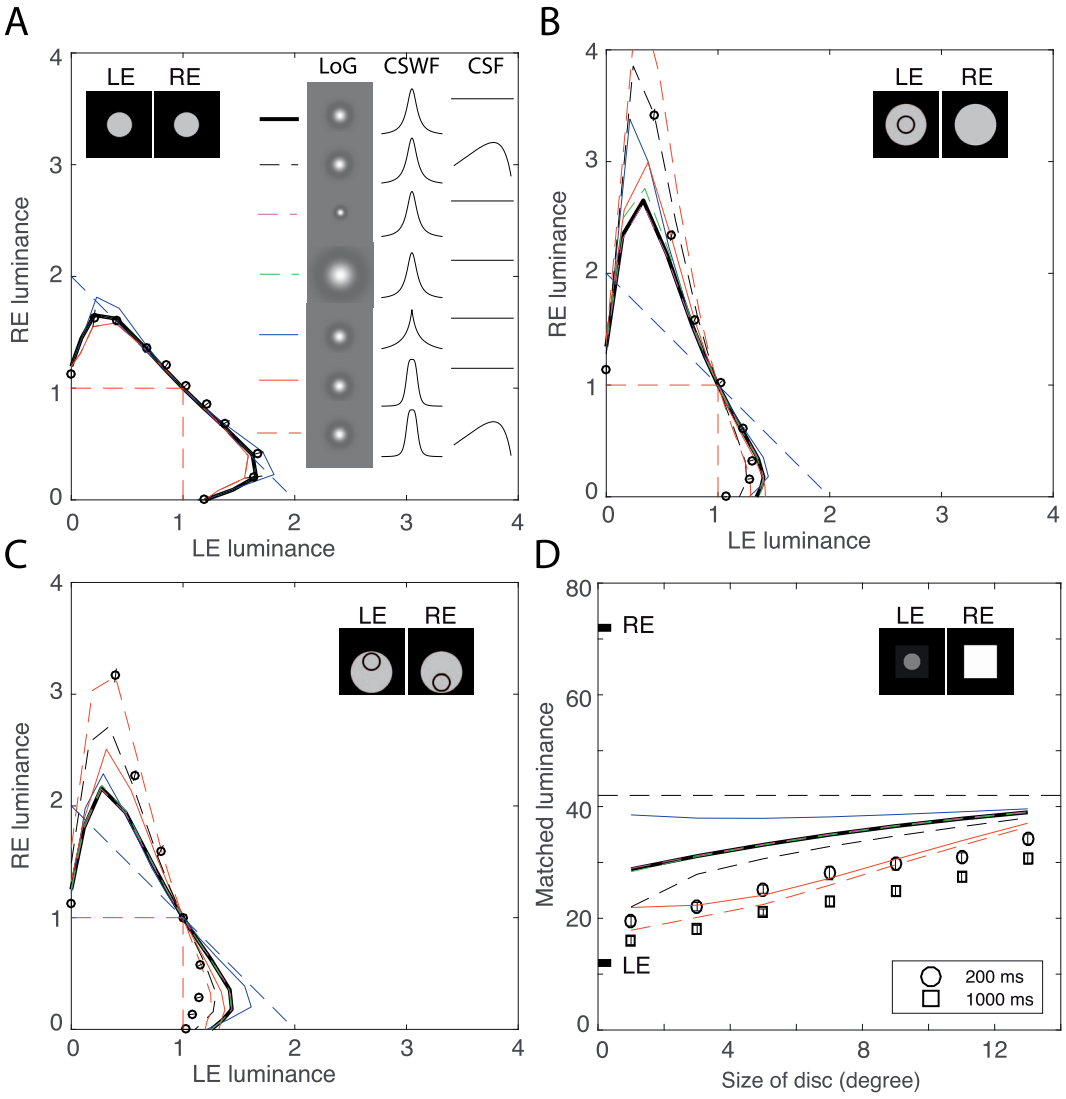

Figure 5. Binocular luminance summation of luminance discs with asymmetric contour information. (A) A sample from Experiment 1 of equal-luminance increment contour when the two eyes view two discs without concentric circles, luminance increment (8 cd/m^2) from a dark background (0.2 cd/m^2). The thick black curve indicates the model prediction using the same LoG filter, contrast space weight function (CSWF), and flat CSF as in Experiment 1. Dashed black curve (overlapped with the thick black curve) indicates the prediction using the same LoG filter and CSWF as in Experiment 1, but using non-flat CSF in the model. Dashed magenta and green curves (also overlapped with the thick black curve) indicate the model predictions using a smaller or a larger LoG filter than that in Experiment 1, while using the same CSWF and flat CSF. Solid blue and red curves indicates the prediction using the same LoG filter and flat CSF but a different CSWF with a slower (blue) or more rapid (red) dropping rate in space weight when going away from the fovea. The dashed red curve (overlapped with the solid red curve) indicates the prediction using a rapid-dropping CSWF and a non-flat CSF while the LoG filter is the same as in Experiment 1. (B) Equal-luminance increment contour when the two eyes view two discs with a concentric circle only in one eye (e.g., LE) (Experiment 2a). The codes for model prediction curves are the same as in (A)(the dashed magenta curve is overlapped with the thick black curve). (C) Equal-luminance increment contour when the two eyes view two discs with circles at different locations of the disc in the two eyes. One eye's circle (e.g., LE's) is in the upper and the other eye's circle (e.g., RE's) is in the lower half of the disc, and the luminance inside the upper circle is the target luminance on which the brightness match was performed (Experiment 2b). The codes for model prediction curves are the same as in (A) (the dashed magenta and green curves are overlapped with the thick black curve). (D) The disc size effect on binocular luminance summation. One eye (e.g., RE) views the luminance, fixed at 72 cd/m^2, of a large square of 14° x 14° on the dark background (0.2 cd/m^2), and the other eye (e.g., LE) views a central disc of variable size (1°, 3°, 5°, 7°, 9°, 11°, 13° in diameters), with luminance fixed at 12 cd/m^2, plus a surrounding square of 14° × 14°, with luminance fixed at 3.7 cd/m^2, on the dark background (0.2 cd/m^2). The luminance of the reference (same incremental luminance in both eyes) was adjusted to match the brightness of the test disc region. The binocularly matched luminance of the central disc was measured and plotted as a function of the disc diameter. The codes for model prediction curves are the same as in (A) (the dashed magenta and green curves are overlapped with the thick black curve).

10), and therefore its weight in binocular summation should increase. To address this model prediction, following Levelt (1965), we assessed binocularly perceived luminance by a matching method when one eye (e.g., RE) views a large $14° \times 14°$ luminance square (72 cd/m^2) and the other eye (e.g., LE) views a luminance disc (12 cd/m^2) with its size varying from 1°–13° in diameter. The luminance of the reference (same incremental luminance in both eyes) was adjusted to match the brightness of the test disc region. Figure 5D shows the matched luminance as a function of disc size. The short black bars indicate the two eyes luminance levels, and the horizontal dashed black line indicates their average (=42 cd/m^2) at which the two eyes have equal weight in the binocular summation. If the matched luminance is below the average line, the LE has more weight in the summation; if above the average line, the RE has more weight in the summation. All experimental data were below the average line; the disc looks dimmed in binocular view, even though the other eye is viewing a very bright large square. The smaller the disc, the more dimmed it appears in binocular view, consistent with Levelt (1965) and confirming the model prediction. We tested two stimulus durations. When binocularly viewing with a long duration (1000 ms), the LE's disc luminance has more weight in the binocular summation than at a short duration (200 ms). Again, installing a nonflat CSF provides better model predictions (dashed black and red curves) than flat CSF, whereas models with different LoG filters give similar predictions. The model prediction is much improved by increasing the dropping rate of the space weight in CSWF when going away from the fovea.

Experiment 3: Binocular combination of second-order stimuli

To test whether the model can also predict second-order binocular combination by simply adding a rectification in the signal path of each eye, we conducted four experiments to evaluate binocular combination of contrast-modulation (CM) gratings when the carrier contrast (CC) was constant or varied.

Experiment 3a: Binocular phase combination of CM gratings with constant carrier contrast

Figure 6A shows the results of binocular CM phase combination when the CC was fixed at 0.2 and the CM phases in the two eyes were -45° and 45°, respectively. When the CM depth varied in the two eyes, no matter whether the two eyes carriers were correlated, uncorrelated or anticorrelated, the perceived CM phase (colored markers in Figure 6A) followed the prediction of linear summation (overlapped with solid colored curves), consistent with (Zhou et al., 2014). All of our five nested models correctly predict the results (solid colored curves) because they all predict equal weights of the two eyes in binocular combination when the CC was identical in the two eyes.

Experiment 3b: Binocular phase combination of CM gratings with varied carrier contrast

To test the models when the two eyes have unequal weights, we ran the experiment of binocular CM phase combination when the interocular CC ratio varied, but the CM depth was identical in the two eyes (=0.7; Figure 6B). The results (colored markers) were similar to those of first-order stimuli (see Figure 8B), deviating from the prediction of linear summation (black dashed curve), no matter whether the two eyes' carriers are correlated, uncorrelated, or anticorrelated, which were correctly predicted by the full model (Figure 2A; colored curved).

Experiment 3c: Binocular CM depth combination of CM gratings with constant carrier contrast

To further test our models, we measured the CM depth by matching a monocular reference CM depth to the test one. When the CC was fixed (=0.2) in the two eyes, and the CM depth varied in the RE and was fixed (=0.7) in the LE (Figure 6C), the monocular reference CM depth (always in the LE; CC = 0.2 and CM varied) was matched to the test pattern. The matched CM depth also followed linear summation, consistent with the previous study (Zhou et al., 2014).

Experiment 3d: Binocular CM depth combination of CM gratings with varied carrier contrast

In this experiment, the CC varied (0–0.2) in the RE and fixed (=0.2) in the LE, and the CM depth was identical in the two eyes (=0.7). Similar to contrast constancy in binocular contrast combination of first-order stimuli, the binocular matched CM depth also remains constant (≈0.7) even when the CC varied in one eye (Figure 6D). The monocular reference CM grating (CC = 0.2 and CM varied) was always in the LE.

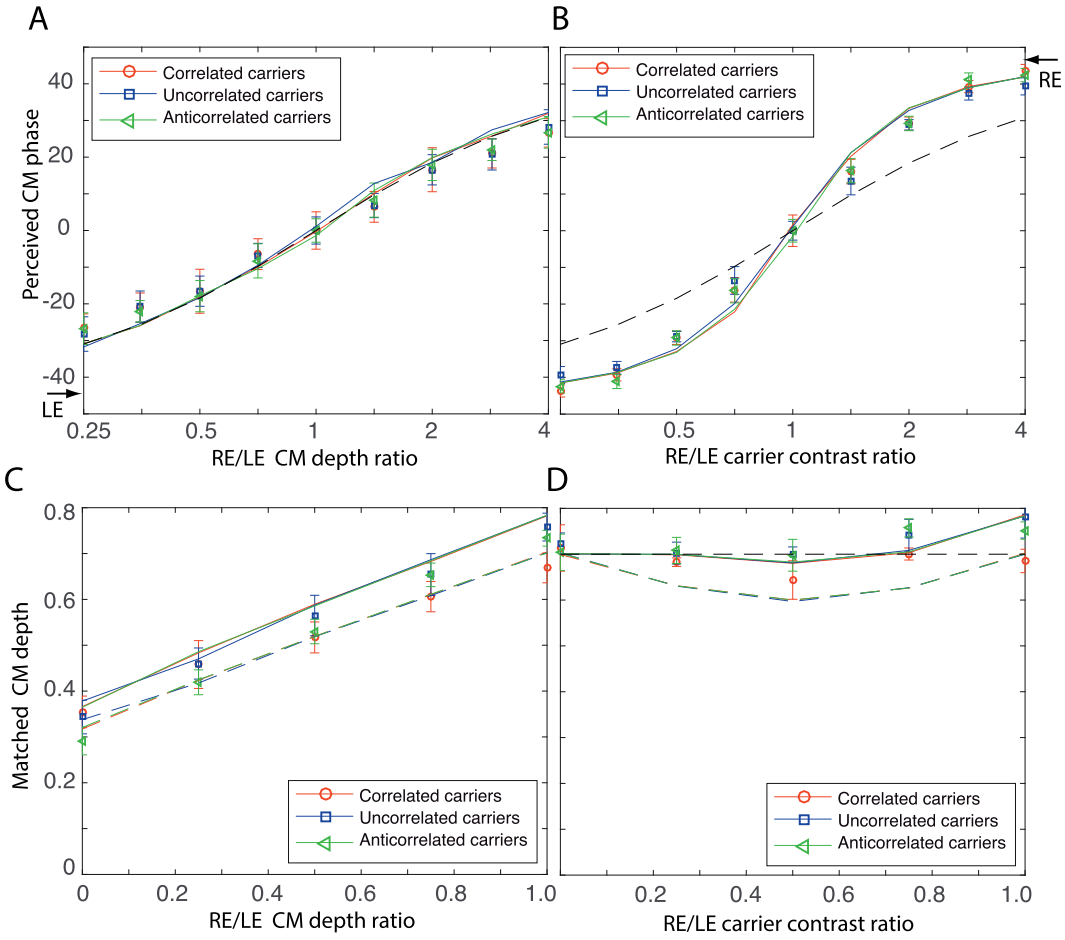

Figure 6. Binocular combination of second-order sinewave gratings either when interocular contrast modulation (CM) depth ratio varied (base CM depth = 0.7) at fixed carrier contrast (CC = 0.2) or when interocular CC ratio varied (base CC = 0.2) at fixed CM depth (=0 .7) in the two eyes. Smooth solid curves are predictions of the full model (Figure 2A) with parameters in Table 4. (A) Binocular CM phase combination when interocular CM depth ratio varied at fixed CC. The black dashed curve indicates the prediction of linear summation (overlapped with solid colored curves). (B) Binocular phase combination when interocular CC ratio varied at fixed CM depth. The black dashed curve indicates the prediction of linear summation. (C) Binocular CM depth combination when interocular CM depth ratio varied at fixed CC. The colored dashed curves indicate the predictions of the first nested model (contrast-weighted summation model). (D) Binocular CM depth combination when interocular CC ratio varied at fixed CM depth. The colored dashed curves indicate the predictions of the first nested model (contrast-weighted summation model). The black dashed horizontal line indicates CM depth constancy in binocular combination of second order stimuli even when the CC varied in one eye.

Modeling

F test for comparison of nested models

Let N_p be the number of model parameters and N_{data} be the number of observed data points. We have the number of degrees of freedom $v = N_{data} - N_p$, and the reduced chi-square is given by $\chi^2_v = \chi^2/v$. If Model a is nested within Model b, the F test that tests whether Model b significantly improves data fitting is given by,

$$F_{a,b} = \frac{\frac{\chi^2(a)-\chi^2(b)}{v(a)-v(b)}}{\frac{\chi^2(b)}{v(b)}}, \quad (26)$$

which compares the variance between models a and b with the variance inside model b and has F distribution with $[v(a) - v(b), v(b)]$ degree of freedom. When the F value is large enough, Model a can be rejected at a small false-rejection probability p(F).

The AIC for comparison of different models

We used the Akaike information criterion (AIC), a measure of the relative goodness of fit of a statistical model developed by Akaike (1974), to compare different models. Let N_p be the number of model parameters and L_{Max} be the maximized value of the likelihood function for the estimated model, AIC is

			JS		PZ		KK		DM		Average				
	Np	*v*	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	*F* test	*p*(F)	AICc
Mod1	**4**	370	1411.1	3.82	776.1	2.10	1814.9	4.92	1272.0	3.45	2600.2	7.03			2608.3
Mod2	**5**	369	878.0	2.39	697.7	1.90	1378.9	3.75	1227.7	3.34	1978.5	5.36	116	<0.001	1988.7
Mod3	**6**	368	477.4	1.30	662.2	1.80	725.1	1.98	477.5	1.30	664.0	1.80	729	<0.001	676.2
Mod4	**7**	367	477.2	1.30	662.2	1.81	713.0	1.94	476.3	1.30	664.0	1.81	0	1	678.3
DSKL	**8**	366	413.6	1.13	658.7	1.80	677.0	1.85	397.2	1.09	575.1	1.57	57	<0.001	591.5

Table 1. Fitting statistics of five nested binocular combination models. *Np*: the number of parameters; *v*: degrees of freedom.

defined as $AIC = 2N_p - 2\ln L_{\text{Max}}$. Assuming that the errors are normally distributed and independent, after ignoring the constant term, AIC is given by

$$AIC = \chi^2 + 2N_p. \quad (27)$$

To give a greater penalty for additional parameters, Burnham and Anderson (2002) recommended the AIC with a correction for finite sample sizes (AICc), which is given by,

$$AICc = AIC + \frac{2N_p(N_p+1)}{N_{\text{data}} - N_p - 1}, \quad (28)$$

where N_{data} is the number of observed data points.

Test five nested binocular combination models

With compression of logarithmic luminance function and including both TCE and TLE in gain-control, we tested five nested binocular combination models (Model 1: Equation 16; Model 2: Equation 19; Model 3: Equation 20; Model 4: Equation 21; DSKL: Equation 23).

Table 1 shows chi square values for model fitting and statistical comparisons of Models 1–5, in which a previous model is nested within its successor. Without gain-control of gain-enhancement, i.e., $\beta = 0$, the DSKL model is simplified to be Model 4, and without gain-enhancement, i.e., the gain-enhancement threshold $g_e = \infty$, Model 4 is further simplified to be Model 3 (the DS model with asymmetric double-layer interocular gain-controls). When the double gain-controls are symmetric, i.e., $\alpha = 1$, Model 3 is simplified to be Model 2 (the DS model with symmetric double-layer interocular gain-controls), and when the gain-control threshold $g_c = 0$ or TCExTLE $\gg 1$, Model 2 is further simplified to be Model 1 (contrast-and-luminance-weighted summation model). The comparison of two neighboring models was made through an *F* test with the *F* value given in the row of the second and subsequent models (*F* test and its *p* value are only shown for the average data). Except for the step modifying Model 3 to Model 4, each step of model modification achieved significant improvement in data fitting; the previous model could be rejected with a very small (<0.001) probability of false rejection. However, without gain control of the gain enhancement (Model 5: DSKL model), the gain enhancement itself in Model 4 failed to further improve the data fitting for all observers.

Test of luminance compressors

We tested three luminance compressors by inserting each of them before the DSKL model: (1) Stevens' law; (2) monocular luminance gain control; (3) luminance logarithmic function. When luminance gain-control threshold = 0 or when luminance $\gg$ threshold, monocular luminance gain-control is simplified to be Stevens' law. Table 2a compares the three nested models: (1) the model without a luminance compressor (or with a linear luminance compressor); (2) the model with Stevens' luminance compressor; and (3) the model with monocular luminance gain-control (LG). Table 2b compares two nested models: (1) the model without a luminance compressor (or with a linear luminance compressor); and (2) the model with a luminance logarithmic function (LOG). The *F* tests show that including a luminance compressor significantly improves model fitting. Upgrading the compressor from Steven's law to luminance gain-control further improves model fitting significantly. The models with LG and LOG compressors have similar fitting performance. Their reduced chi-squares and AICc are very close, in both average and individual data.

Test of luminance energy

Our previous studies (Ding & Levi, 2014) showed that the mean luminance influenced interocular suppression. Reducing the mean luminance in one eye reduced its suppression of the other eye. Indeed, including the total luminance energy (TLE) in the interocular gain-control significantly improved the fitting performance of the DSKL+LOG model (Table 3: noTLE vs. withTLE-in-GC). However, including the TLE in the gain-enhancement path did not improve the

			JS		PZ		KK		DM		Average				
	Np	v	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	*F* test	*p*(F)	AICc
–	**6**	368	602.2	1.64	787.2	2.14	1277.1	3.47	523.5	1.42	1183.3	3.22			1195.5
Steven	**7**	367	560.1	1.53	756.3	2.06	1147.0	3.13	500.1	1.36	1048.4	2.86	47	<0.001	1062.7
LG	**9**	365	410.6	1.12	656.0	1.80	683.3	1.87	390.1	1.07	563.6	1.54	157	<0.001	582.1

Table 2a. Fitting statistics of models with or without luminance compressors. *Np*: the number of parameters; v: degrees of freedom.

			JS		PZ		KK		DM		Average				
	Np	v	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	*F* test	*p*(F)	AICc
–	**6**	368	602.2	1.64	787.2	2.14	1277.1	3.47	523.5	1.42	1183.3	3.22			1195.5
LOG	**8**	366	413.6	1.13	658.7	1.80	677.0	1.85	397.2	1.09	575.1	1.57	193	<0.001	591.5

Table 2b. Fitting statistics of models with or without luminance compressors. *Np*: the number of parameters; v: degrees of freedom.

fitting performance (noTLE vs. withTLE-in-GE), and including the TLE in both the interocular gain-control and gain-enhancement (withTLE-in-GCGE) made the fitting performance worse than that of only including the TLE in the gain-control path (withTLE-in-GC). Although we do not have direct evidence for how the luminance affects interocular enhancement, the fitting statistics in Table 3 suggests that the TLE might have no little or effect on interocular enhancement.

Model parameters

Table 4 shows the parameters of the DS (Model 3) and DSKL models combining with a suitable luminance compressor, either a luminance gain-control (LG) or a logarithmic function (LOG).

Model simulation

The DSKL model was originally developed for the combination of binocular signals within a relatively narrow spatial-frequency band on a fixed luminance background, with an influence of stimuli (mask) in other spatial frequency bands on this combination (Ding et al., 2013b; Ding & Sperling, 2006, 2007). The gain-control was dependent on both the signal and the mask. The total contrast energy (TCE), summed over all spatial-frequency channels, was used in the gain-control paths. In the current study, although the model was proposed for a more general case, it was only tested partially. In Experiment 1, partial model parameters (Table 4) were estimated with perceived luminance of a dichoptic luminance disc, but others were just selected to be reasonable, because we did not have sufficient data to fit them. To further test the model, we ran Experiments 2 and 3 with more complex stimuli and used the model obtained from Experiment 1, with a different combination of LoG filter, space-weight function, and CSF, to predict their results. We also used the model to predict the results of our previous studies. Theoretically, with model parameters (Table 4) from fitting and a selected space-weight function (e.g., Equation 10 with $R_0 = 0.1°$ and $R_{max} = 22.6°$) and Laplacian of Gaussian (LoG) filter (e.g., Equation 7 with $\sigma_i = 0.045°$), we can use the model to predict binocular combination of any two images. However, we confined our simulations to two similar inputs to avoid binocular rivalry, which has a different mechanism that is out of scope of our model predictions.

Let $I_L(x, y)$ and $I_R(x, y)$ be two inputs to the two eyes, the output, $\hat{I}(x, y)$, of the model (Figure 2A) is given by Equation 23. A MATLAB program extracted the image features (e.g., phase, luminance, contrast, and contrast modulation) from both input and output profiles. First, we need to test if the simulation can predict the data of Experiment 1. After selecting a suitable value of b_i, the TCE defined in Equation 8 can

			JS		PZ		KK		DM		Average				
	Np	v	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	*F* test	*p*(F)	AICc
noTLE	**7**	367	482.2	1.31	727.1	1.98	1068.8	2.91	450.7	1.23	871.8	2.38			886.1
withTLE in GC	**8**	366	413.6	1.13	658.7	1.80	677.0	1.85	397.2	1.09	575.1	1.57	188.8	<0.001	591.5
withTLE in GE	**8**	366	482.2	1.32	727.1	1.99	1068.8	2.92	450.7	1.23	871.8	2.38	–	–	886.1
withTLE in GCGE	**8**	366	434.1	1.19	661.1	1.81	687.6	1.88	399.8	1.09	612.3	1.67	–	–	628.7

Table 3. Fitting statistics of models with or without including luminance energy. *Np*: the number of parameters; v: degrees of freedom.

Models	Luminance compressors			TLE	Binocular combination models				
	Z	p	q	η	g_c	α	γ	g_e/g_c	β
DS+LG	2.76 ± 0.24	1.79 ± 0.14	1.47 ± 0.11	0.64 ± 0.07	0.007 ± 0.001	1.10 ± 0.004	2.30 ± 0.11	∞	0
DS+LOG	1.55 ± 0.18	2.04 ± 0.16	–	0.65 ± 0.07	0.009 ± 0.002	1.10 ± 0.004	2.32 ± 0.11	∞	0
DSKL+LG	3.08 ± 0.25	1.99 ± 0.12	1.56 ± 0.09	0.32 ± 0.04	0.011 ± 0.003	0.78 ± 0.06	1.93 ± 0.04	2.40 ± 0.40	0.61 ± 0.06
DSKL+LOG	2.29 ± 0.26	1.74 ± 0.09	–	0.39 ± 0.04	0.011 ± 0.002	0.82 ± 0.06	1.94 ± 0.05	2.56 ± 0.43	0.64 ± 0.07

Table 4. Model parameters.

be calculated from the edge contrast (Equation 14). Similarly, after selecting a suitable value of k, the total luminance energy (TLE) defined in Equation 9 is equal to the average luminance energy of the disc defined in Equation 13. With these suitable values of b_i and k , the simulation provides accurate predictions of the data in Figures 3 and 4, as shown by the solid black curves in each panel.

Binocular luminance summation of luminance discs with asymmetric contour information

In Experiment 1, for mathematical simplicity, we assumed a flat CSF for the modeling. Here, we installed a non-flat CSF for model simulation. Because of suprathreshold stimuli used in this study, we used the CSF obtained from a judging task (Mannos & Sakrison, 1974), given by,

$$CSF = 2.6(0.0192 + 0.114f_s)\exp\left(-(0.114f_s)^{1.1}\right), \quad (29)$$

which has a peak of value 1.0 at $f_s = 8.0$ cpd and a zero-frequency intercept of 0.05. Figure 7A shows its response to a luminance disc with or without a concentric circle. At each black-white border, a sharp increase-and-then-decrease can be observed. In the simulation, the binocular luminance output is defined as the mean luminance of the central area, 0.3° away from a border (e.g., between two vertical dashed lines in Figure 7A), where the luminance response is almost flat. When luminance discs in the two eyes have no concentric circle (Experiment 1), there is little difference in model predictions between flat and nonflat CSFs (see sample predictions in Figure 5A). However, when viewing two asymmetric stimuli (Experiment 2: Figure 5B–D), although the model with a flat CSF correctly predicts the direction of the asymmetry, the prediction shows less asymmetry than the data. Installing the nonflat CSF further increases the asymmetry in prediction when viewing asymmetric stimuli, which much improves the model performance.

We tested three LoG filters, whose peak spatial frequencies occur at 20 cpd (small), 10 cpd (middle), and 5 cpd (large), respectively, in the simulation to extract local contrast. Figure 7B shows their local contrast outputs for a luminance disc with a concentric circle. After selecting suitable constant b_i in Equation 8 to make the TCE to be equal to the edge contrast for each LoG filter, the models with different LoG filters are very similar in predictions of the results of Experiments 1 and 2.

We also tested three space weight functions, given by Equations 10, 24, and 25, which have different dropping rates in weight when going away from the

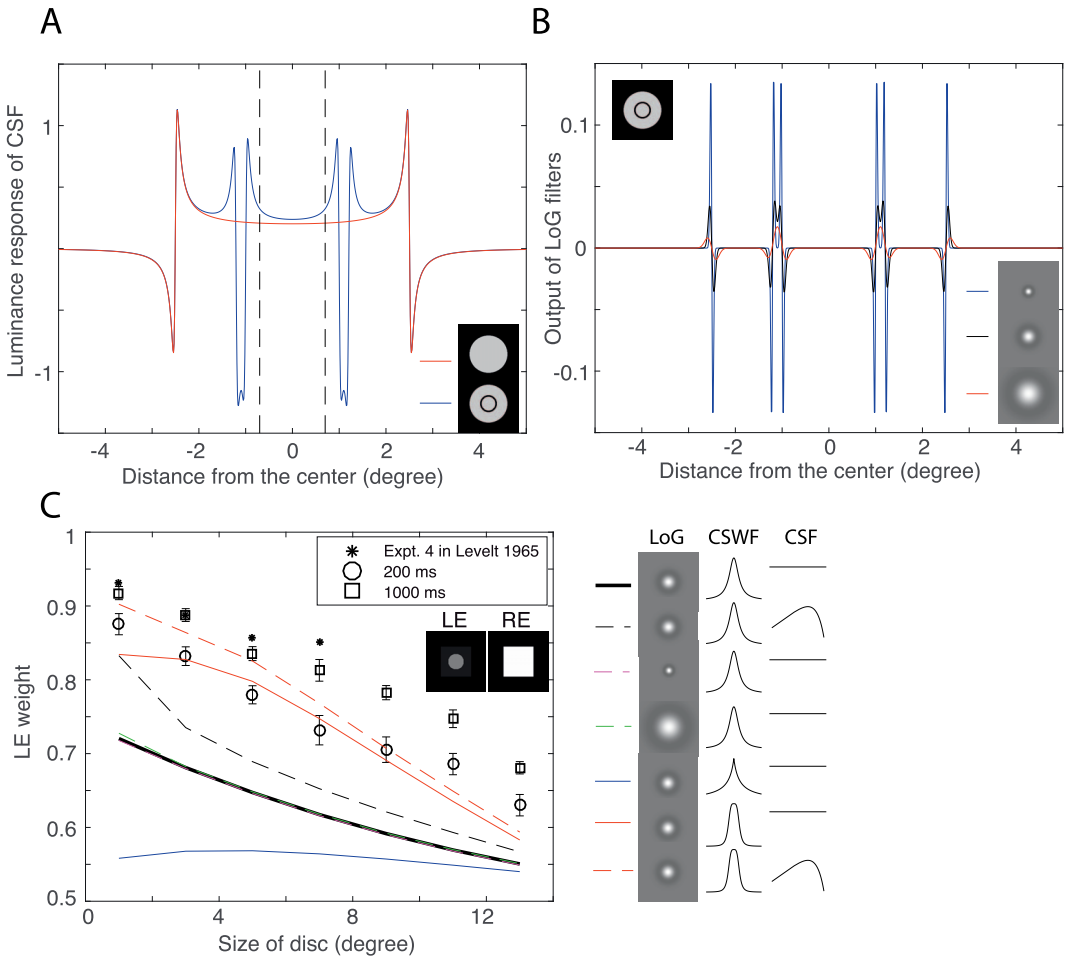

Figure 7. Model simulations of binocular luminance summation of luminance discs with asymmetric contour information. (A) Luminance response of a nonflat CSF following luminance compression for a luminance disc (8 cd/m^2 luminance increment from dark background) with (blue) or without (red) a concentric circle. (B) Local contrast output of three LoG filters following luminance compression for a luminance disc (8 cd/m^2 luminance increment from dark background) with a concentric circle. (C) The disc size effect on binocular luminance summation: LE weight in luminance summation as a function of disc size (replotted from Figure 5D). The RE views the luminance, fixed at 72 cd/m^2, of a square of 14° × 14° on the dark background (0.2 cd/m^2), and the LE views a central disc of variable size (1°, 3°, 5°, 7°, 9°, 11°, 13° in diameters), with luminance fixed at 12 cd/m^2, plus a surrounding square of 14° × 14°, with luminance fixed at 3.7 cd/m^2, on the dark background (0.2 cd/m^2). The sum of the two eyes' weights equal one. The star markers show the results from Levelt (1965). Both simulation and data show that the LE has more weight in the luminance summation, and that the LE's weight decreases when the disc size increases.

fovea. When the stimulus is shifted from the fovea, we assumed that the space weight is proportional to $1/r$ (slow CSWF given by Equation 24), $1/r^2$ (middle CSWF given by Equation 10), or $1/r^4$ (fast CSWF given by Equation 25). When a luminance disc increases in size, its edge length increases (proportional to its radius r), which would fully compensate the contrast weight decrease for the slow CSWF. Therefore the model with the slow CSWF would predict equal weighting of the two eyes in luminance summation for discs with different sizes (blue curve in Figure 7C). In contrast, the model with the middle or fast CSWF predicts lower weight for the eye with a larger size disc, because the increasing edge length is not sufficient to compensate the decrease in weight (Figure 7C). Following Levelt (1965), we assumed that the binocularly perceived luminance is a linear weighted sum of the two eyes' luminances, *i.e.*, $B = w_L L + w_R R$ with $w_L + w_R = 1$, which is valid for most middle part of equal-luminance increment contour in Figures 5A–C. From data of Experiment 2c (Figure 5D), we calculated the LE's weight when viewing different size of discs and replotted them in Figure 7C. The LE with a small disc dominates binocular summation when the RE views a large luminance square, and the LE's weight decreases when the disc size increases, thus rejecting the slow CSWF. The middle CSWF was assumed based on the fact that the cone density is proportional to $1/r^2$

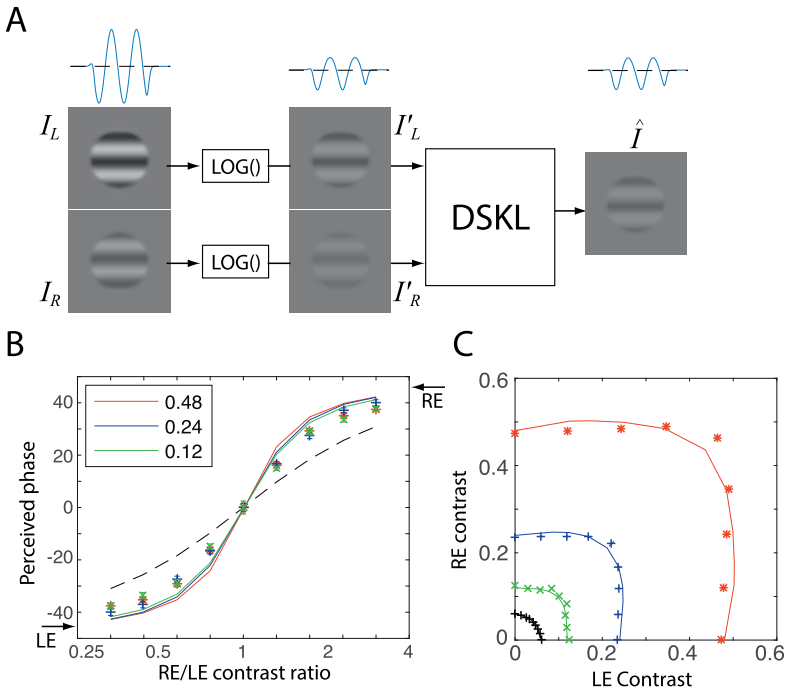

Figure 8. Model simulation: Binocular combination of first-order sinewave gratings. The best-fit model parameters of Figures 3 and 4 (Table 4) were used in the simulation. (A) Two input sinewave gratings, their compressed images by a luminance compressor (logarithmic luminance function), and the binocular output of the DSKL model. The one-dimensional sine waves are also shown for the LE input, its compressed image and the binocular output. (B) Simulation of binocular phase combination. The two eye's input phases are −45° and 45°, respectively as indicated by arrows on the sides, and the base contrast is 0.48 (red), 0.24 (blue), or 0.12 (green). The dashed black curve represents the prediction of the linear summation model. (C) Simulation of binocular contrast combination: equal contrast contours. The two eye input phases are identical and the reference contrast is 0.48 (red), 0.24 (blue), 0.12 (green), or 0.06 (black). Experimental data from Ding and Levi (2016) are indicated as colored markers.

(Curcio et al., 1987; Curcio et al., 1990; Perry & Cowey, 1985; Williams, 1988). However, its dropping rate is still not sufficient to explain the eye asymmetry in Figure 7C. After installing the fast CSWF, the model performance was much improved. Apparently, beyond the retina, the central visual system also contributes to the fovea-periphery asymmetry, e.g., the fovea receives more attention than the periphery (Levi, Klein, & Aitsebaomo, 1985).

To compare with previous studies, we also replotted the results of Experiment 4 from Levelt (1965; stars in Figure 7C), which are more asymmetric than our results of short stimulus duration (200 ms; circles in Figure 7C), but comparable with those of long stimulus duration (1000 ms; squares in Figure 7C), most likely because he used an adjustment task that has a long stimulus duration.

Binocular combination of luminance modulated (first-order) gratings

In the following, we simulated our previous binocular phase and contrast combination experiments (Ding & Levi, 2016). The input first-order gratings (Figure 8A) are given by,

$$I_{\mathrm{L}}(x,y) = I_0(1 + m_{\mathrm{L}} \sin(2\pi f_s y + \theta_{\mathrm{L}}) h_{\mathrm{cir}}(x,y))$$

and

$$I_{\mathrm{R}}(x,y) = I_0(1 + m_{\mathrm{R}} \sin(2\pi f_s y + \theta_{\mathrm{R}}) h_{\mathrm{cir}}(x,y)), \quad (30)$$

where $h_{\mathrm{cir}}(x,y)$ is a circular (radius = 1.5°) space window function (1 inside and 0 outside the circle) with a blurred edge (Gaussian envelope: sigma = 0.1°).

As shown in Figure 8A, the two input images are first compressed by a luminance compressor (logarithmic luminance function), and then go through the DSKL model to be combined. The nonlinear luminance compression distorts the input luminance profile, resulting in the black-white asymmetry—the amplitude in the black direction is larger than in white direction (see 1-D sine waves in Figure 8A), consistent with previous studies (He & MacLeod, 1998; Lu & Sperling, 2012).

For simulation of binocular phase combination, $\theta_{\mathrm{L}} = -45°$, and $\theta_{\mathrm{R}} = 45°$, at one base contrast (=max[m_{L}, m_{R}]) and contrast ratio $m_{\mathrm{R}}/m_{\mathrm{L}}$, the binocular output grating $\hat{I}(x,y)$ was first calculated from

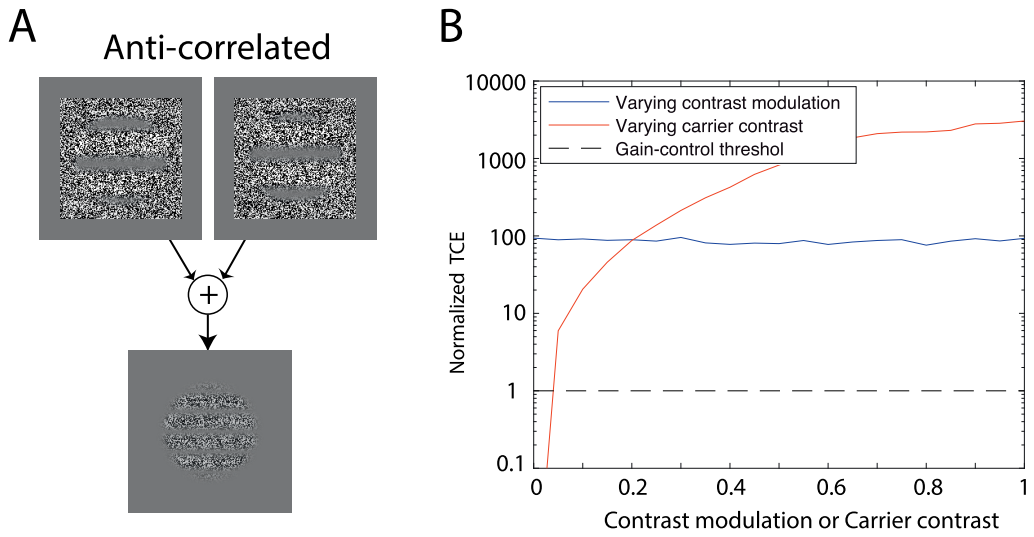

Figure 9. (A). Linear summation of two eyes' CM gratings with 90° out of phase when their carriers are anti-correlated. (B). Total contrast energy (TCE) normalized with the gain-control threshold when the contrast modulation (CM) depth is varied but the carrier contrast (CC) is fixed at 0.2 (blue) or when the CM depth is fixed at 0.7 but the CC is varied (red). The gain-control threshold is also indicated (black dashed line).

Equation 23, and then, after Fourier transformation, its phase (the perceived phase) was estimated as the phase at spatial frequency f_s (the peak frequency in the spatial frequency domain). Figure 8B shows that the predicted perceived phase as a function of contrast ratio at one base contrast (colored curves) were a good fit to the data (colored markers) of Ding and Levi (2016). To simulate binocular contrast combination, the two eyes have identical input phase, i.e., $\theta_L = \theta_R$. Figure 8C shows equal contrast contours when the interocular contrast ratio varies at one base contrast. The model output contrast is estimated by (max − min) / (max + min) of the output image $\hat{I}(x, y)$. As shown in Figure 8C, the model gives good predictions of the binocular contrast combination data of Ding and Levi (2016).

Binocular combination of contrast modulated (second-order) gratings

Similarly, we can simulate the binocular combination of second-order contrast-modulated (CM) gratings. The input CM gratings are given by,

$$I_L(x,y) = I_0(1 + (1 + CM_L \sin(2\pi f_s y + \theta_L)h_{cir}(x,y)) \times N_L(x,y)h_{sq}(x,y))$$

and

$$I_R(x,y) = I_0(1 + (1 + CM_R \sin(2\pi f_s y + \theta_R)h_{cir}(x,y)) \times N_R(x,y)h_{sq}(x,y)), \quad (31)$$

where $\theta_L = -45°$ and $\theta_R = 45°$, CM_L and CM_R are contrast modulation depth, $N_L(x, y)$ and $N_R(x, y)$ are the carriers (binary noise), $h_{cir}(x, y)$ is a circular space window, the same with the one in Equation 30, and $h_{sq}(x, y)$ is a square window (4.5° × 4.5°) with sharp edges.

When the carriers in the two eyes are positively correlated or uncorrelated, linear summation of the two eyes' CM gratings with different phases has the same spatial frequency f_s of CM as the monocular inputs and the combined phase of CM gratings can be measured at f_s. However, when the two eyes have CM gratings of different phases with anticorrelated carriers, their linear combination doubles the spatial frequency ($2f_s$)(Figure 9A) because the background contrast is canceled in the combination. This is not consistent with experimental observations (Zhou et al., 2014)—the binocular combined CM grating has the same spatial frequency of CM as the monocular inputs, no matter whether the carriers in the two eyes are correlated, uncorrelated, or anticorrelated. Therefore, in the binocular combination model of second-order (CM) stimuli, we added a rectification in the signal path of each eye before the binocular combination site (Figure 2A) to avoid the cancelation when the two eyes' carriers are anti-correlated. However, in the standard FRF (filter-rectify-filter) structure for second-order stimuli, the second filter can be either before or after the binocular combination (our simulation cannot tell the difference).

The binocular combination model of second-order stimuli is almost the same as that of first-order stimuli except the final output is given by

$$\hat{I}(x,y) = |\hat{I}_L(x,y)| + |\hat{I}_R(x,y)|.$$

In the simulation, we first Fourier transformed the

binocular output $\hat{I}(x, y)$ and then measured the phase and the amplitude (the binocular CM depth) at the spatial frequency f_s (the peak frequency in the spatial frequency domain), which is equivalent to recovering the CM sinewave using the second filter of the FRF model after the binocular combination site. In the same way, two monocular CM depths were measured, and the higher one was used to normalize the binocular CM depth.

By simulation, we found that the TCE is independent of the contrast modulation (CM) depth, but dependent on the carrier contrast (CC; Figure 9B). When CC = 0.2, the TCE is about 90 times higher than the gain-control threshold at all CM depths (blue curve in Figure 9B). When the CC is identical in the two eyes, their weights in binocular combination are equal even though the interocular CM depth ratio varied, resulting in linear binocular phase summation, no matter whether the carrier is correlated, noncorrelated, or anticorrelated (color curves in Figure 6A). Consistent with (Zhou et al., 2014), the results (colored markers in Figure 6A) of our Experiment 3a when the CC was fixed at 0.2 but the interocular CM ratio varied, bore out this prediction.

The aforementioned simulation suggests that the interocular interaction is not based on local contrast, which varies in a sinusoidal fashion for a CM grating, but depends on the total contrast energy (TCE) summed over space, which remains constant even when the local contrast varies. This results in equal weighting of the two eyes in the binocular combination of CM gratings when their CCs are equal. However, it is unclear whether the TCE is summed over the whole or just a part of a retina. Further work is needed to answer this question, and to reveal more detail of the space weighting function for TCE and TLE.

When TCE $\gg$ threshold and is equal in the two eyes, all five nested models (see Model section) have the property that the two eyes have equal weight in the binocular combination. With this property, all five models predict "linear" summation of CM gratings with identical CC in the two eyes (Georgeson & Schofield, 2016; Zhou et al., 2014), although the models themselves are nonlinear in general. In other words, the data in Figure 6A cannot distinguish among models—any model with a symmetric structure in the two eyes, can predict them. They cannot even test any model parameters, which could be any value without affecting the model prediction of linear summation in Figure 6A.

Therefore, we ran Experiment 3b (Figure 6B) to vary the CC in the two eyes to test whether a binocular model, originally proposed for first-order stimuli, also works for second-order stimuli. Again, all five nested models gave similar predictions (solid colored curves in Figure 6B), consistent with our previous study on first order stimuli (Ding et al., 2013b), in which we also found that the phase data alone were not enough to tell the difference among our five nested models—the amplitude information was needed to test the models.

Although all five nested models gave the two eyes equal weights in the binocular combination of CM gratings with identical CC, the weights given by different models might be different. For example, the weight given by the first nested model (contrast weighted summation model) was 0.5, while the weight given by the full model (the DSKL model) was around 0.56, no matter whether the two eyes' carriers were correlated, uncorrelated, or anticorrelated. The slightly higher weight (~0.56) predicted by the DSKL model is due to the interocular enhancement, which is included in the DSKL model but absent in the contrast-weighted model. However, in this study, we did not have sufficient data to test this small difference between the two models' predictions (see Figure 6C).

The contrast-weighted summation model proposed by Zhou et al. (2014) is similar to our first nested model in predictions of the results of Experiments 3a–3d. Georgeson and Schofield (2016) elaborated the Zhou model by adding two monocular channels in parallel to the binocular one. The final output was assumed to be the maximum of the three channels. Its monocular channels play a key role in preventing cancellation at the antiphase disparity (180°), providing more reasonable predictions of CM detection data when the CM gratings were 180° out of phase in the two eyes (Georgeson & Schofield, 2016). However, when CM phase disparity is 0°–90°, its binocular channel always wins in the maximum operation, making the elaborated model identical to Zhou et al.'s model. Therefore both Zhou et al.'s model and its elaboration are similar to our first nested model in predictions of the results of Experiments 3a–3d.

Rebalancing binocular vision in amblyopia

The motivation for including the luminance energy in the model comes from our previous study on rebalancing binocular vision in amblyopia (Ding & Levi, 2014). Individuals with amblyopia suffer imbalanced binocular vision. When they open both eyes, the dominant eye (DE) strongly suppresses the nondominance eye (NDE), rendering it almost blind. To reduce the DE's suppression, we placed a neutral density (ND) filter in front of the DE to reduce its mean luminance, and we successfully rebalanced their binocular vision with their two eyes having equal contribution to binocular combination using the appropriate ND filter. Here, we run model simulations to see how the model works for binocular rebalancing in amblyopia.

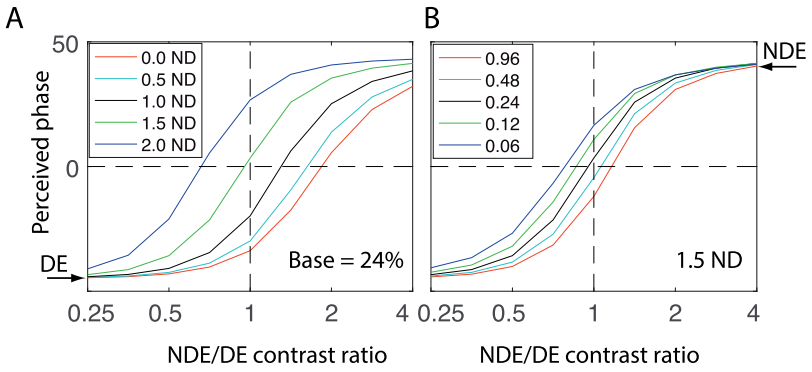

Figure 10. Simulation of rebalancing binocular vision in amblyopia (Observer: GD; Ding & Levi, 2014). The two eyes were presented with two luminance gratings that differ in phase (DE = −45° and NDE = 45° as indicated on the two sides of the plot) and contrast. The mean luminance in the NDE was always 26.2 cd/m^2 (without ND filter), but the mean luminance in the DE varied by placing an ND filter before it.

To simulate amblyopic binocular combination, we used our previous model and its parameters (Ding et al., 2013a; without luminance compressors) but using the product of total contrast energy and total luminance energy (TCE × TLE) for calculating interocular gain-control. Figure 10 shows the simulation of binocular phase combination for amblyopic observer GD with a ND filter in front of her DE (Ding & Levi, 2014). The two eyes were presented with two luminance gratings that differed in phase and contrast. The mean luminance in the NDE was always 26.2 cd/m^2 (without an ND filter), but the mean luminance in the DE was varied by placing a ND filter before it. The TCE was proportional to the stimulus contrast, i.e., TCE $\sim C^{\gamma}$, and the TLE was proportional to the mean luminance, i.e., TLE $\sim \mathrm{Lum}^{\eta}$ (η value comes from Table 4). We normalized the TLE so that TLE =1 with no ND filter (i.e., Lum = 26.2 cd/m^2), because the previous model was proposed for experiments performed without varying the mean luminance and with the assumption of TLE = 1. Figure 10A shows the model prediction at base contrast = 24% when an ND filter is placed before the DE. Without an ND filter before the DE (ND = 0: red curve in Figure 10A), binocular vision is very asymmetric. When both eyes have equal contrast (contrast ratio = 1, dashed vertical line), the perceived phase is almost the same as that of the DE, reflecting the fact that the NDE makes very little contribution to the percept. Placing ND filters, from 0.0–2.0 log units, before the DE, results in the perceived phase curve systematically shifting up and to the left (from red to blue curves) towards the balance point (perceived phase = 0 when the contrast ratio = 1). When a 1.5 ND filter was placed before the DE and the base contrast varied from 96% to 6% (Figure 10B), the perceived phase curve also systematically shifted, demonstrating that the binocular balance in amblyopia also depends on the base contrast. These predictions are consistent with our previous findings (Ding & Levi, 2014).

Discussion

Binocular combination is a complex process. Simple models (like binocular averaging, power summation or vector summation) might succeed for one task, but they often fail to explain others. Our goal was to develop a unified model that can explain all (or at least many) binocular phenomena—the current 2D model represents a first step towards building such a model. Although only some of its parameters were determined by luminance summation data, our simulations show that the model with the same parameters can also explain binocular phase and contrast combination of first-order gratings, binocular CM phase and depth combination of second-order CM gratings, and binocular luminance summation with asymmetric contour information in the two eyes. Installing a luminance compressor enables the model to operate over a large range of luminance input, and its properties at high luminance levels are consistent with Stevens' power law for brightness perception (see the following). Including luminance energy in the model represents an important development from our previous models, significantly improving the model fits to our current luminance data, and also providing a reasonable account for rebalancing binocular vision in amblyopia by reducing the luminance input in the DE (using an ND filter). Although interocular enhancement could not be observed directly in this study because of strong interocular suppression, including it in the model significantly improved its performance in interpreting binocular luminance combination data, and is consistent with our previous study for binocular phase and contrast combination (Ding et al., 2013b).

Comparison with other models

We tested five nested binocular combination models and found that when combined with a luminance compressor and including luminance energy in inter-ocular gain-control, both the Ding–Sperling (Model 3) and the DSKL (Model 5) models provide accurate predictions of binocular luminance summation of both luminance increments and decrements over a large range of luminance input. Do other binocular combination models give similar results if they are combined with a luminance compressor? To answer this question, we need to install a binocular contrast combination model to explain luminance data because most models were developed for contrast tasks. For this purpose, we assumed that the luminance increment/decrement of a target from its background is perceived through luminance contrast perception. Figure 11 shows a binocular contrast combination model combined with a luminance compressor (LOG). The binocular contrast combination model can be either the power summation model (Legge, 1984), the normalization model (Moradi & Heeger, 2009), the first stage of the two-stage model (Meese et al., 2006), or the first stage of two-stage + Maximum operator model (Georgeson et al., 2016).

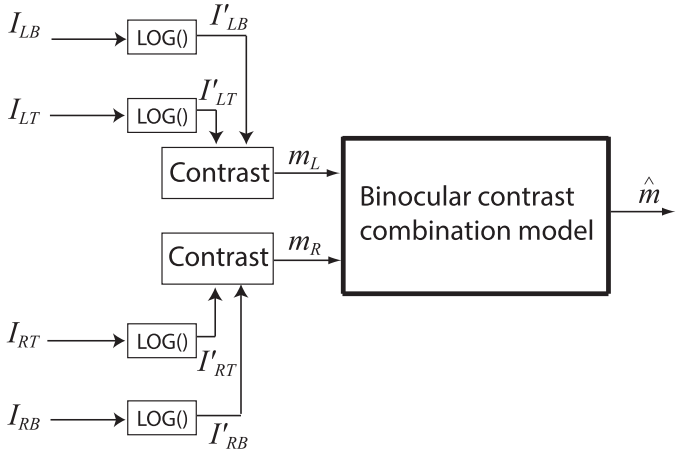

Figure 11. A binocular contrast combination model is modified to predict binocular luminance summation. The luminance inputs of target and background first go through a luminance compressor (LOG) (Equation 4), and then two eyes' luminance contrasts (edge contrast: Equation 15) are extracted and binocularly combined. The binocular contrast combination model can be either the power summation model (Legge, 1984), the normalization model (Moradi & Heeger, 2009), the first stage of the two-stage model (Meese et al., 2006) or the first stage of the two-stage + Maximum operator model (Georgeson et al., 2016).

Power summation model

Legge (1984) proposed a square summation model to explain binocular contrast summation, which was later extended to the power summation model:

$$\widehat{m} = (m_{\mathrm{L}}^{\gamma} + m_{\mathrm{R}}^{\gamma})^{1/\gamma}, \quad (32)$$

where m_{L} and m_{R}, the edge contract of the two eyes' discs, are given by Equation 15.

Contrast normalization model

Moradi and Heeger (2009) proposed the contrast normalization model to explain their fMRI data in binocular combination. The contrast is first extracted from both eyes and pooled together, and then the pooled contrast exerts gain control to the two eyes separately before binocular combination. The binocularly-combined contrast is given by:

$$\widehat{m} = \frac{m_{\mathrm{L}}^{\gamma}}{S^{\gamma} + m_{\mathrm{L}}^{\gamma} + m_{\mathrm{R}}^{\gamma}} + \frac{m_{\mathrm{R}}^{\gamma}}{S^{\gamma} + m_{\mathrm{L}}^{\gamma} + m_{\mathrm{R}}^{\gamma}}. \quad (33)$$

The two-stage model

The two-stage model was originally proposed to account for monocular, dichoptic, half binocular, and binocular contrast discrimination (Meese et al., 2006), and was later used to account for the perceived contrast of a cyclopean sinewave (Baker et al., 2007). Its first stage is for binocular contrast combination and its second stage, located after binocular combination site, is for contrast discrimination. The output of the first stage of the two-stage model is given by,

$$\widehat{m} = \frac{1}{S + m_{\mathrm{L}} + m_{\mathrm{R}}} m_{\mathrm{L}}^{\gamma} + \frac{1}{S + m_{\mathrm{L}} + m_{\mathrm{R}}} m_{\mathrm{R}}^{\gamma}, \quad (34)$$

Two-stage + Max model (GWMB model)

Recently, the two-stage model has been expanded to account for a total of 11 different types of dipper function with various combinations of pedestal, increments and decrements for targets (Georgeson et al., 2016). In order to explain half binocular contrast discrimination with decrements for targets, they extended the model by including two independent monocular channels. The final perceived contrast is assumed to be the maximum contrast of the three channels (one binocular and two monocular channels). The contrast in the binocular channel $\widehat{m}_{\mathrm{B}}$ is also given by Equation 34, and the contrast in two monocular channels is given by,

			JS		PZ		KK		DM		Average		
	Np	v	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	χ^2	χ^2/v	AICc
Legge	**3**	371	908.6	2.45	2129.8	5.74	1501.3	4.05	786.6	2.12	2042.0	5.50	2048.1
Norm	**4**	370	908.6	2.46	2129.8	5.76	1501.3	4.06	786.6	2.13	2042.0	5.52	2050.2
Two-stage	**4**	370	591.7	1.60	698.1	1.89	1069.5	2.89	583.8	1.58	938.1	2.54	946.2
GWMB	**4**	370	698.9	1.89	1879.6	5.08	1119.7	3.03	666.6	1.80	1472.0	3.98	1480.1
DS	**6**	368	477.4	1.30	662.2	1.80	725.1	1.98	477.5	1.30	664.0	1.80	676.2
DSKL	**8**	366	413.6	1.13	658.7	1.80	677.0	1.85	397.2	1.09	575.1	1.57	591.5

Table 5. Model comparison (combining with a LOG compressor). *Np*: the number of parameters; v: degrees of freedom; DS model: Equation 20 (Model 3); DSKL model: Equation 23 (Model 5); Two-stage model: Equation 31; Two-stage + Max model (GWMB model): Equations 31-33; Legge Model (power summation model): Equation 29; contrast normalization model: Equation 30.

$$\widehat{m}_{\mathrm{L}} = \frac{1}{S + m_{\mathrm{L}}} m_{\mathrm{L}}^{\gamma}, \quad \widehat{m}_{\mathrm{R}} = \frac{1}{S + m_{\mathrm{R}}} m_{\mathrm{R}}^{\gamma}. \quad (35)$$

The final perceived contrast is given by

$$\widehat{m} = \max[\widehat{m}_{\mathrm{L}}, \widehat{m}_{\mathrm{R}}, \widehat{m}_{\mathrm{B}}], \quad (36)$$

Without adding any new parameters, the modified model (Equations 34–36) had better performance in predicting 11 different types of dipper function than the original one of Equation 34 (Georgeson et al., 2016). Unlike the independent monocular channels (no effect from the other eye) assumed in Equations 34–36, a monocular channel in the binocular combination model proposed by Georgeson and Schofield (2016) for second-order stimuli receives suppression from the other eye.

Table 5 shows the fitting statistics for model comparison when fitting the results of Experiment 1. Figure 12 only demonstrates the fits of these models for luminance increment contours on a dark background (0.2 cd/m^2). Interestingly, the first stage of the two-stage model also predicts both Fechner's paradox and the winner-take-all features after combining it with a LOG luminance compressor, providing good predictions for all the luminance data of Experiment 1. However, the first stage of the two-stage+Max model (the GWMB model) performed poorly; it failed to predict Fechner's paradox in the luminance increment contour (Figure 12) because the Max operator in Equation 36 kills the paradox. The power summation (Legge model) and the contrast normalization models also failed to predict Fechner's paradox. The prediction from the contrast normalization model was almost identical to that of the Legge model, and therefore is not shown in Figure 12.

Luminance compressors

Figure 13 shows the outputs of luminance compressors in a combined model with the DS or DSKL models (parameters from Table 4). When Lum $>$ 5 cd/m^2, in both DS and DSKL combined models, the compressors' output have slopes around 0.33–0.5, consistent with the exponents in Steven's power law for a 5-degree target and a point source respectively based on brightness perception experiments (Stevens, 1960). Both luminance compressors LG and LOG have similar features: The luminance output increases rapidly initially when the input increases from dark, and then becomes compressed at a high luminance level. Combined with interocular gain control, this behavior provides a unified explanation of both Fechner's paradox and the winner-take-all phenomenon.

(1) When one eye's (e.g., LE's) luminance increases from a dark background, both the compressed luminance (the compressor's output) and the edge contrast (Equation 15) increases rapidly (increasing slope $>$ 1), and therefore its suppression of the other eye (e.g., RE) increases rapidly (the RE's output decreases more rapidly than the LE's increase), resulting in Fechner's paradox.

(2) When one eye's (e.g., LE's) luminance increases from a light background, the output-increasing slope is flat (0.33–0.5). Its suppression of the other eye can be canceled by its own increase in the binocular combination, resulting in the winner-take-all phenomenon. With the same amplitude of input change, the decrement has more contrast energy than the increment because of luminance compression, i.e., the black-white asymmetry (Lu & Sperling, 2012; MacLeod, 1972). The model correctly predicts that the winner-take-all phenomenon will be more evident in the equal-luminance decrement contour than in the equal-luminance increment contour (Figures 3 and 4).

He and MacLeod (1998) applied contrast-modulation flicker to study the dynamics and spatial resolution of the light adaptation process. Because of the black-white asymmetry, their observers saw brightness changes (contrast-modulation flicker) when the contrast of a high frequency grating was modulated periodically between zero and a peak value, even though its spatial and temporal average luminance was kept constant. They found that the light adaptation

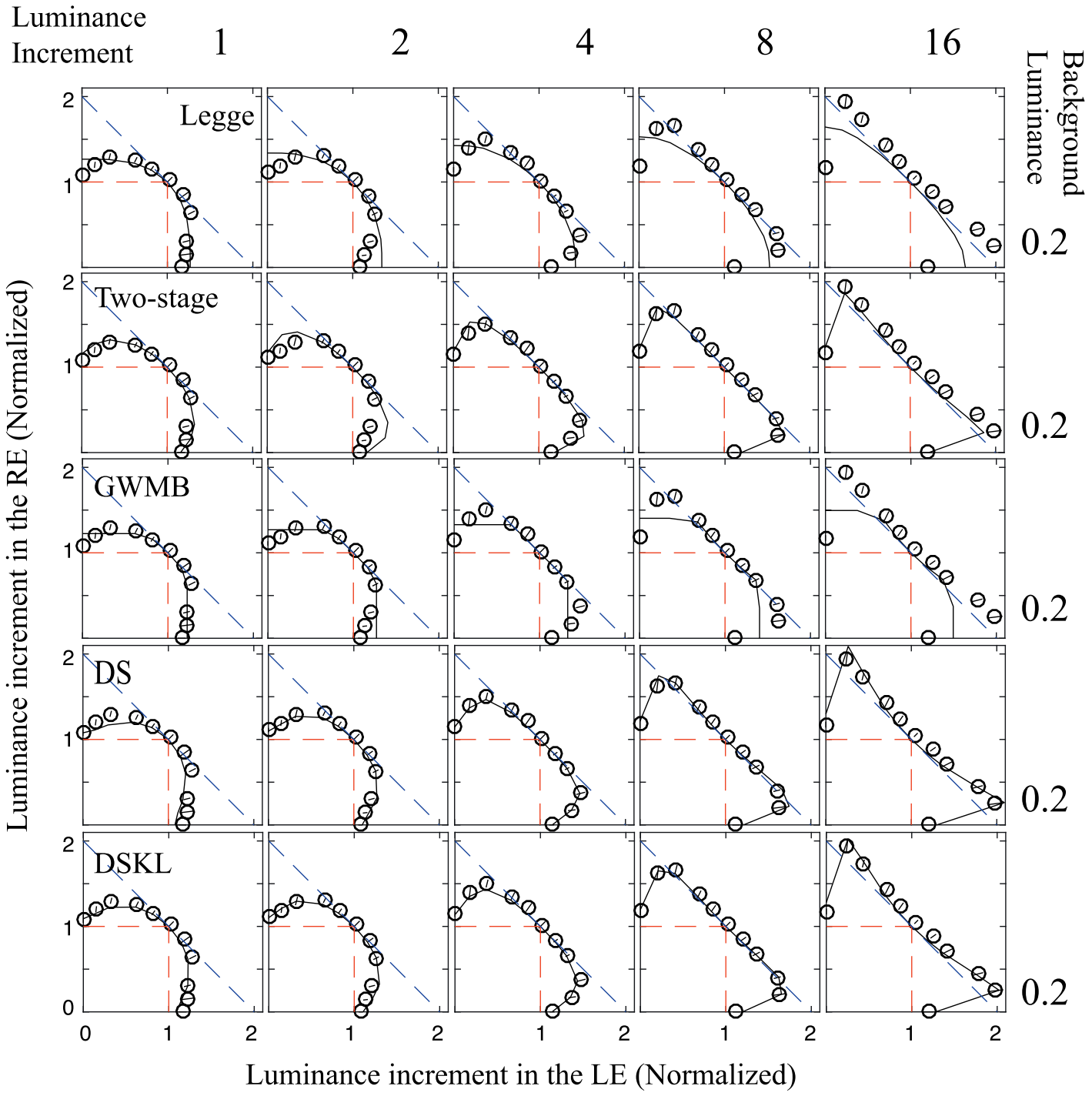

Figure 12. Prediction of equal luminance increment contours at dark background of five binocular combination models combining with a luminance compressor (LOG). For the two-stage and the GWMB models, only the first stage was tested. The contrast normalization model gave the identical prediction to the Legge model (power summation model).

acted very rapidly (< 20 ms) with very high spatial resolution (> 100 cpd), at which the grating itself was too fine to be visually resolved as a pattern. They argued that the light adaptation occurs very early and very locally, either residing within individual photoreceptors or operating on signals from individual receptors.

The site of interocular interaction

The lateral geniculate nucleus (LGN) is the first location at which interocular interactions occur in the retino-geniculo-cortical pathways. Although each lamina of the LGN receives direct retinal afferents from only one eye (the dominant eye), stimuli in the other

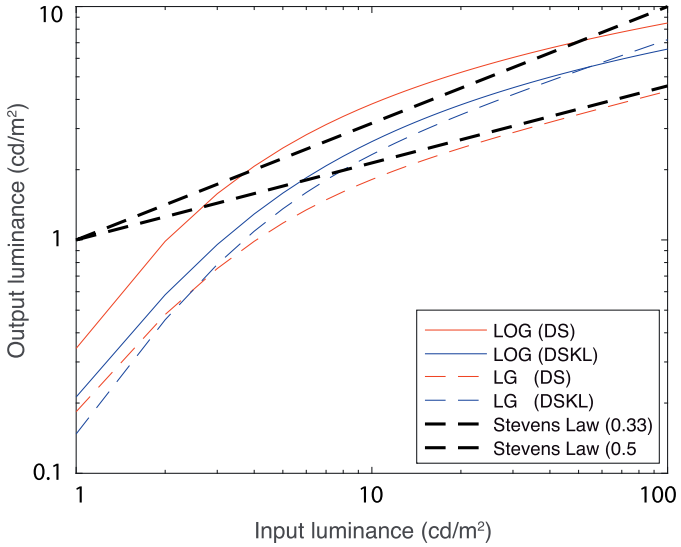

Figure 13. Simulation of luminance compressors. The model parameters were given in Table 4. The Stevens' power law, based on brightness perception experiments (Stevens, 1960), was also shown for a 5° target (power = 0.33) or a point source (power = 0.5) in the dark.

eye (the nondominant eye) also inhibit (nondominant inhibition) or facilitate (nondominant facilitation) its dominant response to the dominant-eye stimuli.

Nondominant eye inhibition has been well documented in both cat and monkey LGN (Guido, Tumosa, & Spear, 1989; Moore, Spear, Kim, & Xue, 1992; Pape & Eysel, 1986; Rodieck & Dreher, 1979; Sanderson, Bishop, & Darian-Smith, 1971; Schroeder, Tenke, Arezzo, & Vaughan, 1990; Sengpiel, Blakemore, & Harrad, 1995; Singer, 1970; Tong, Guido, Tumosa, Spear, & Heidenreich, 1992; Wang, Dreher, & Burke, 1994; Xue, Ramoa, Carney, & Freeman, 1987). Sengpiel et al. (1995) reported that the mean dominant response in cat LGN was inhibited by up to 44% during stimulation through the nondominant eye with gratings of any orientation; This nondominant inhibition was essentially independent of the orientation and direction of drift of the gratings shown to the nondominant eye. Even an iso-oriented grating presented to the nondominant eye exerted inhibition on the dominant eye's response, which was essentially independent of interocular phase differences (Sengpiel et al., 1995), in contrast to the phase-selective (e.g., disparity selective) interaction seen in many binocularly driven cells in visual cortex (Ohzawa & Freeman, 1986; Ohzawa, Sclar, & Freeman, 1985). This nondominant inhibition in the LGN may provide the physiological substrate for the interocular gain-control in binocular combination, because (1) cells in the LGN generally receive excitatory input from only one eye, but their dominant responses are influenced by the other eye; (2) the interocular gain-control is also essentially independent of the orientation and direction of drift of the gratings (Ding & Sperling, 2006, 2007); and (3) the interocular gain-control is also essentially independent of interocular phase differences of iso-oriented gratings (Ding & Sperling, 2006, 2007).

More interestingly, the literature on the cat dLGN (dorsal LGN) also supports our assumption of the gain-control of gain-control (e.g., the RE's gain-control to the LE itself is gain-controlled by the LE, blue path in Figure 2A), which was initially introduced to the model to avoid excessive interocular-suppression in binocular combination (Ding & Sperling, 2006, 2007). Pape and Eysel (1986) found that afferent activity from the dominant eye reduced inhibition from the nondominant eye in the cat dLGN. They demonstrated this by a localized destruction of the receptive field area in the dominant eye that greatly increased the nondominant eye inhibition. Wang et al. (1994) also demonstrated the strong dependence of nondominant inhibition on the maintained discharge rate of the cat dLGN cell by quantifying the nondominant inhibition as the ratio of the inhibition dip to the maintained discharge rate (the "dip ratio"). They found that, at low maintained discharge rates, the dip ratio approached unity; as the maintained discharge rate increased, the dip ratio decreased.

The origin of the nondominant inhibition of the LGN cells remains uncertain. It could conceivably be due to direct inhibitory interaction between right-eye and left-eye laminae of the LGN, via inhibitory interneurons (Guillery, 1966), which is supported by the fact that the visual cortex removal does not reduce nondominant influences on LGN cells (Pape & Eysel, 1986; Sanderson et al., 1971; Singer, 1970; Tumosa, McCall, Guido, & Spear, 1989). However, Ding and Sperling (2006, 2007) found that, although essentially orientation independent, the interocular gain control was somewhat more effective for vertical and horizontal gratings than diagonal ones, which suggests that some of the gain control is likely of cortical origin, i.e., arises beyond the LGN, because the LGN cells are essentially indifferent to orientation.

Unlike nondominant eye inhibition, nondominant eye facilitation in the LGN is less well documented. Schroeder et al. (1990) found nondominant eye facilitation in monkey LGN using flash-evoked multiunit activity (MUA) record. They reported that the amplitude of the binocular response is more than twice that of the monocular dominant response, and is not simply the sum of the monocular dominant and nondominant responses. This nondominant facilitation might provide the physiological substance for the interocular gain-enhancement in binocular combination. Although the model assumes gain-control of gain-enhancement (e.g., the RE-to-LE's gain-enhancement itself receives the gain-control form the LE; red path in

Figure 2A), it is still unclear whether the afferent activity from the dominant eye reduces the nondominant eye facilitation in the LGN.

Similar to retinal ganglion cells, LGN cells have circular RFs when stimulated through the dominant eye, typically with an antagonistic center-surround organization. However, through the nondominant eye, the LGN cells have no center-surround receptive-field structure and receive similar nondominant eye influence from spots or bars of either light or dark (Guido et al., 1989; Rodieck & Dreher, 1979; Sanderson et al., 1971; Singer, 1970). Thus, the nondominant eye influence in the LGN cells may be nonspecific for input images. This is consistent with our findings that the interocular interaction depends on the total contrast energy (TCE), which is a weighted sum over space and spatial-frequency domains, no matter whether it is a signal or noise. However, the magnitude of nondominant inhibition depended on the spatial frequency of the gratings presented to the nondominant eye (Moore et al., 1992), with its maximum being at a spatial frequency close to the one eliciting the strongest excitation in the dominant eye. Although we also found that the interocular gain-control was spatial-frequency dependent (Ding & Sperling, 2006, 2007), and included a spatial-frequency modulation transfer function (FMTF) into the model (Figure 2B), it is still unclear whether the maximum frequency of FMTF is related to the signal frequency.

Future work

Although the current model is quite complicated, there is still a long way to go in order to build a unified model. Further work is needed to answer the following questions: (1) What are the details of the space weighting functions for calculation of TCE and TLE? Are they different or the same for TCE and TLE? Does the weighted summation take place over the entire retina or just in a localized retinal area? (2) What are the details of the LoG filter for extracting local contrast? How many different LoG filters are required to extract local contrast at different scales? (3) What are the details of the FMTF (spatial-frequency modulation transfer function) for interocular gain-control? Does it depend on the signal frequency? (4) Is it possible to design an experiment to reveal the interocular enhancement in normal vision? Why does the brain need interocular enhancement? (5) How does binocular disparity affect interocular interactions? (6) How does interocular correlation affect interocular interactions? (7) Does binocular combination need a motor/sensory fusion mechanism?

The current model is two-dimensional and static, which might be a specific case of a dynamic 3D model that reaches a steady state at a constant binocular disparity. Although it is still unclear how to include depth cues and dynamic terms into the model to give vivid 3D predictions from two 2D inputs, the current model might be able to serve as an initial step for future development of a more general model.

Keywords: interocular gain-control, interocular gain-enhancement, luminance compressor, luminance energy, contrast modulation

Acknowledgments

This work was supported by National Eye Institute grant R01EY020976 from the National Eye Institute. Publication made possible in part by support from the Berkeley Research Impact Initiative (BRII) sponsored by the UC Berkeley Library.

Commercial relationships: none.
Corresponding author: Jian Ding.
Email: jian.ding@berkeley.edu.
Address: School of Optometry and the Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA, USA.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Anderson, P. A., & Movshon, J. A. (1989). Binocular combination of contrast signals. *Vision Research*, *29*(9), 1115–1132.

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, 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. (2012). Nonlinearities in the binocular combination of luminance and contrast. *Vision Research*, *56*(0), 1–9.

Banton, T., & Levi, D. M. (1991). Binocular summation in vernier acuity. *Journal of the Optical Society of America A*, *8*(4), 673–680.

Bearse, M. A., Jr., & Freeman, R. D. (1994). Binocular summation in orientation discrimination depends

on stimulus contrast and duration. *Vision Research*, *34*(1), 19–29.

Blake, R., & Wilson, H. (2011). Binocular vision. *Vision Research*, *51*(7), 754–770.

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

Burnham, K. P., & Anderson, D. (2002). *Model selection and multi-model inference: A pratical informatio-theoric approch*. New York: Springer-Verlag.

Campbell, F. W., & Green, D. G. (1965). Monocular versus Binocular Visual Acuity. *Nature*, *208*(5006), 191–192.

Cogan, A. I. (1987). Human binocular interaction: Towards a neural model. *Vision Research*, *27*(12), 2125–2139.

Cohn, T. E., & Lasley, D. J. (1976). Binocular vision: two possible central interactions between signals from two eyes. *Science*, *192*(4239), 561–563.

Curcio, C. A., Sloan, K. R., Kalina, R. E., & Hendrickson, A. E. (1990). Human photoreceptor topography. *Journal of Comparative Neurology*, *292*(4), 497–523.

Curcio, C. A., Sloan, K. R., Packer, O., Hendrickson, A. E., & Kalina, R. E. (1987). Distribution of cones in human and monkey retina: Individual variability and radial asymmetry. *Science*, *236*(4801), 579–582.

de Weert, C. M., & Levelt, W. J. M. (1974). Binocular brightness combinations: Additive and nonadditive aspects. *Perception & Psychophysics*, *15*(3), 551–562.

Ding, J., Klein, S. A., & Levi, D. M. (2013a). Binocular combination in abnormal binocular vision. *Journal of Vision*, *13*(2):14, 1–31, 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, doi:10.1167/13.2.13. [PubMed] [Article]

Ding, J., & Levi, D. (2015). Interocular contrast gain control plus monocular luminance gain control can explain binocular luminance summation. *Journal of vision*, *15*(12):263, doi:10.1167/15.12.263. [Abstract]

Ding, J., & Levi, D. M. (2014). Rebalancing binocular vision in amblyopia. *Ophthalmic and Physiological Optics*, *34*(2), 199–213.

Ding, J., & Levi, D. M. (2016). Binocular contrast discrimination needs monocular multiplicative noise. *Journal of Vision*, *16*(5):12, 1–21, doi:10.1167/16.5.12. [PubMed] [Article]

Ding, J., & Sperling, G. (2006). A gain-control theory of binocular combination. *Proceedings of the National Academy of Science 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 Unversity Press.

Engel, G. (1969). The autocorrelation function and binocular brightness mixing. *Vision Research*, *9*(9), 1111–1130.

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

Geisler, W. S. (1981). Effects of bleaching and backgrounds on the flash response of the cone system. *The Journal of Physiology*, *312*, 413.

Geisler, W. S. (1983). Mechanisms of visual sensitivity: backgrounds and early dark adaptation. *Vision Research*, *23*(12), 1423–1432.

Georgeson, M. A., & Schofield, A. J. (2016). Binocular functional architecture for detection of contrast-modulated gratings. *Vision Research*, *128*, 68–82.

Georgeson, M. A., Wallis, S. A., Meese, T. S., & Baker, D. H. (2016). Contrast and lustre: A model that accounts for eleven different forms of contrast discrimination in binocular vision. *Vision Research*, *129*, 98–118.

Grossberg, S., & Kelly, F. (1999). Neural dynamics of binocular brightness perception. *Vision Research*, *39*(22), 3796–3816.

Guido, W., Tumosa, N., & Spear, P. D. (1989). Binocular interactions in the cat's dorsal lateral geniculate nucleus. I. Spatial-frequency analysis of responses of X, Y, and W cells to nondominant-eye stimulation. *Journal of Neurophysiology*, *62*(2), 526–543.

Guillery, R. (1966). A study of Golgi preparations from the dorsal lateral geniculate nucleus of the adult cat. *Journal of Comparative Neurology*, *128*(1), 21–49.

Hayhoe, M., Benimoff, N., & Hood, D. (1987). The time-course of multiplicative and subtractive adaptation process. *Vision Research*, *27*(11), 1981–1996.

Hayhoe, M., Levin, M., & Koshel, R. J. (1992). Subtractive processes in light adaptation. *Vision Research*, *32*(2), 323–333.

He, S., & MacLeod, D. I. (1998). Contrast-modulation flicker: Dynamics and spatial resolution of the light adaptation process. *Vision research*, *38*(7), 985–1000.

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, doi:10.1167/13.8.3. [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.

Legge, G. E. (1981). A power law for contrast discrimination. *Vision Research*, *21*(4), 457–467.

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

Legge, G. E., & Rubin, G. S. (1981). Binocular interactions in suprathreshold contrast perception. *Perception & Psychophysics*, *30*(1), 49–61.

Lehky, S. R. (1983). A model of binocular brightness and binaural loudness perception in humans with general applications to nonlinear summation of sensory inputs. *Biological Cybernetics*, *49*(2), 89–97.

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

Levi, D. M., Klein, S. A., & Aitsebaomo, A. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, *25*(7), 963–977.

Li, X., Lu, Z.-L., Xu, P., Jin, J., & Zhou, Y. (2003). Generating high gray-level resolution monochrome displays with conventional computer graphics cards and color monitors. *Journal of Neuroscience Methods*, *130*(1), 9–18.

Lu, Z.-L., & Sperling, G. (2012). Black–white asymmetry in visual perception. *Journal of Vision*, *12*(10):8, 1–21, doi:10.1167/12.10.8. [PubMed] [Article]

MacLeod, D. I. (1972). The Schrodinger equation in binocular brightness combination. *Perception*, *1*(3), 321–324.

Mannos, J., & Sakrison, D. (1974). The effects of a visual fidelity criterion of the encoding of images. *IEEE transactions on Information Theory*, *20*(4), 525–536.

Mansfield, J. S., & Legge, G. E. (1996). The binocular computation of visual direction. *Vision Research*, *36*(1), 27–41.

Mante, V., Frazor, R. A., Bonin, V., Geisler, W. S., & Carandini, M. (2005). Independence of luminance and contrast in natural scenes and in the early visual system. *Nature Neuroscience*, *8*(12), 1690–1697.

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B: Biological Sciences*, *207*(1167), 187–217.

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, doi:10.1167/6.11.7. [PubMed] [Article]

Moore, R. J., Spear, P. D., Kim, C. B. Y., & Xue, J.-T. (1992). Binocular processing in the cat's dorsal lateral geniculate nucleus III. Spatial frequency, orientation, and direction sensitivity of nondominant-eye influences. *Experimental Brain Research*, *89*(3), 588–598.

Moradi, F., & Heeger, D. J. (2009). Inter-ocular contrast normalization in human visual cortex. *Journal of Vision*, *9*(3):13, 1–22, doi:10.1167/9.3.13. [PubMed] [Article]

Ohzawa, I., & Freeman, R. D. (1986). The binocular organization of complex cells in the cat's visual cortex. *Journal of Neurophysiology*, *56*(1), 243–259.

Ohzawa, I., Sclar, G., & Freeman, R. D. (1985). Contrast gain control in the cat's visual system. *Journal of Neurophysiology*, *54*(3), 651–667.

Pape, H.-C., & Eysel, U. T. (1986). Binocular interactions in the lateral geniculate nucleus of the cat: GABAergic inhibition reduced by dominant afferent activity. *Experimental Brain Research*, *61*(2), 265–271.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Perry, V. H., & Cowey, A. (1985). The ganglion cell and cone distributions in the monkey's retina: Implications for central magnification factors. *Vision Research*, *25*(12), 1795–1810.

Rodieck, R. W., & Dreher, B. (1979). Visual suppression from nondominant eye in the lateral geniculate nucleus: A comparison of cat and monkey. *Experimental Brain Research*, *35*(3), 465–477.

Sanderson, K., Bishop, P., & Darian-Smith, I. (1971). The properties of the binocular receptive fields of lateral geniculate neurons. *Experimental Brain Research*, *13*(2), 178–207.

Schroeder, C., Tenke, C., Arezzo, J., & Vaughan, H. (1990). Binocularity in the lateral geniculate nucleus of the alert macaque. *Brain research*, *521*(1), 303–310.

Sengpiel, F., Blakemore, C., & Harrad, R. (1995). Interocular suppression in the primary visual cortex: a possible neural basis of binocular rivalry. *Vision Research*, *35*(2), 179–195.

Shooner, C., Hallum, L. E., Kumbhani, R. D., García-Marín, V., Kelly, J. G., Majaj, N. J., ... Kiorpes, L.

(2017). Asymmetric dichoptic masking in visual cortex of amblyopic macaque monkeys. *Journal of Neuroscience*, 1760–1717.

Singer, W. (1970). Inhibitory binocular interaction in the lateral geniculate body of the cat. *Brain Research*, *18*(1), 165–170.

Stevens, S. S. (1960). The psychophysics of sensory function. *American Scientist*, *48*(2), 226–253.

Tong, L., Guido, W., Tumosa, N., Spear, P. D., & Heidenreich, S. (1992). Binocular interactions in the cat's dorsal lateral geniculate nucleus, II: Effects on dominant-eye spatial-frequency and contrast processing. *Visual Neuroscience*, *8*(6), 557–566.

Tumosa, N., McCall, M. A., Guido, W., & Spear, P. D. (1989). Responses of lateral geniculate neurons that survive long-term visual cortex damage in kittens and adult cats. *Journal of Neuroscience*, *9*(1), 280–298.

Wang, C., Dreher, B., & Burke, W. (1994). Non-dominant suppression in the dorsal lateral geniculate nucleus of the cat: laminar differences and class specificity. *Experimental Brain Research*, *97*(3), 451–465.

Williams, D. R. (1988). Topography of the foveal cone mosaic in the living human eye. *Vision Research*, *28*(3), 433–454.

Xue, J., Ramoa, A., Carney, T., & Freeman, R. (1987). Binocular interaction in the dorsal lateral geniculate nucleus of the cat. *Experimental Brain Research*, *68*(2), 305–310.

Yeh, T., Lee, B. B., & Kremers, J. (1996). The time course of adaptation in macaque retinal ganglion cells. *Vision Research*, *36*(7), 913–931.

Yehezkel, O., Ding, J., Sterkin, A., Polat, U., & Levi, D. (2016). Binocular combination of stimulus orientation. *Royal Society Open Science*, *3*(11), 160534.

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, doi:10.1167/14.13.24. [PubMed] [Article]