

Matching and correlation computations in stereoscopic depth perception

Takahiro Doi

Graduate School of Frontier Biosciences, Osaka University,
Toyonaka, Osaka, Japan



Seiji Tanabe

Graduate School of Frontier Biosciences, Osaka University,
Toyonaka, Osaka, Japan



Ichiro Fujita

Graduate School of Frontier Biosciences, Osaka University,
Toyonaka, Osaka, Japan



A fundamental task of the visual system is to infer depth by using binocular disparity. To encode binocular disparity, the visual cortex performs two distinct computations: one detects matched patterns in paired images (matching computation); the other constructs the cross-correlation between the images (correlation computation). How the two computations are used in stereoscopic perception is unclear. We dissociated their contributions in near/far discrimination by varying the magnitude of the disparity across separate sessions. For small disparity (0.03°), subjects performed at chance level to a binocularly opposite-contrast (anti-correlated) random-dot stereogram (RDS) but improved their performance with the proportion of contrast-matched (correlated) dots. For large disparity (0.48°), the direction of perceived depth reversed with an anti-correlated RDS relative to that for a correlated one. Neither reversed nor normal depth was perceived when anti-correlation was applied to half of the dots. We explain the decision process as a weighted average of the two computations, with the relative weight of the correlation computation increasing with the disparity magnitude. We conclude that matching computation dominates fine depth perception, while both computations contribute to coarser depth perception. Thus, stereoscopic depth perception recruits different computations depending on the disparity magnitude.

Keywords: stereo vision, binocular vision, random-dot stereograms, fine stereopsis, coarse stereopsis

Citation: Doi, T., Tanabe, S., & Fujita, I. (2011). Matching and correlation computations in stereoscopic depth perception. *Journal of Vision*, 11(3):1, 1–16, <http://www.journalofvision.org/content/11/3/1>, doi:10.1167/11.3.1.

Introduction

Physiological studies suggest that neural computations encoding binocular disparity are distinct between the ventral and dorsal pathways of the primate visual cortex (Neri, 2005; Orban, Janssen, & Vogels, 2006; Parker, 2007; Tanabe, Umeda, & Fujita, 2004). The distinction is evident in the disparity-tuning function when a neuron responds to an anti-correlated stereogram. Such stereograms eliminate matched patterns between the two eyes by contrast reversing one of the two images that are projected onto the eyes (Julesz, 1971). Along the dorsal pathway, the tuning function is inverted relative to the tuning function obtained from a correlated stereogram (Krug, Cumming, & Parker, 2004; Takemura, Inoue, Kawano, Quaia, & Miles, 2001). This suggests a correlation computation in which the disparity signal (i.e., peak value minus the baseline in the disparity-tuning function) is proportional to the stimulus correlation (i.e., peak value minus the baseline in the cross-correlation function) between the two eyes such that it resembles a cross-correlation of binocular images (Cumming & Parker, 1997). On the other hand, along the ventral pathway,

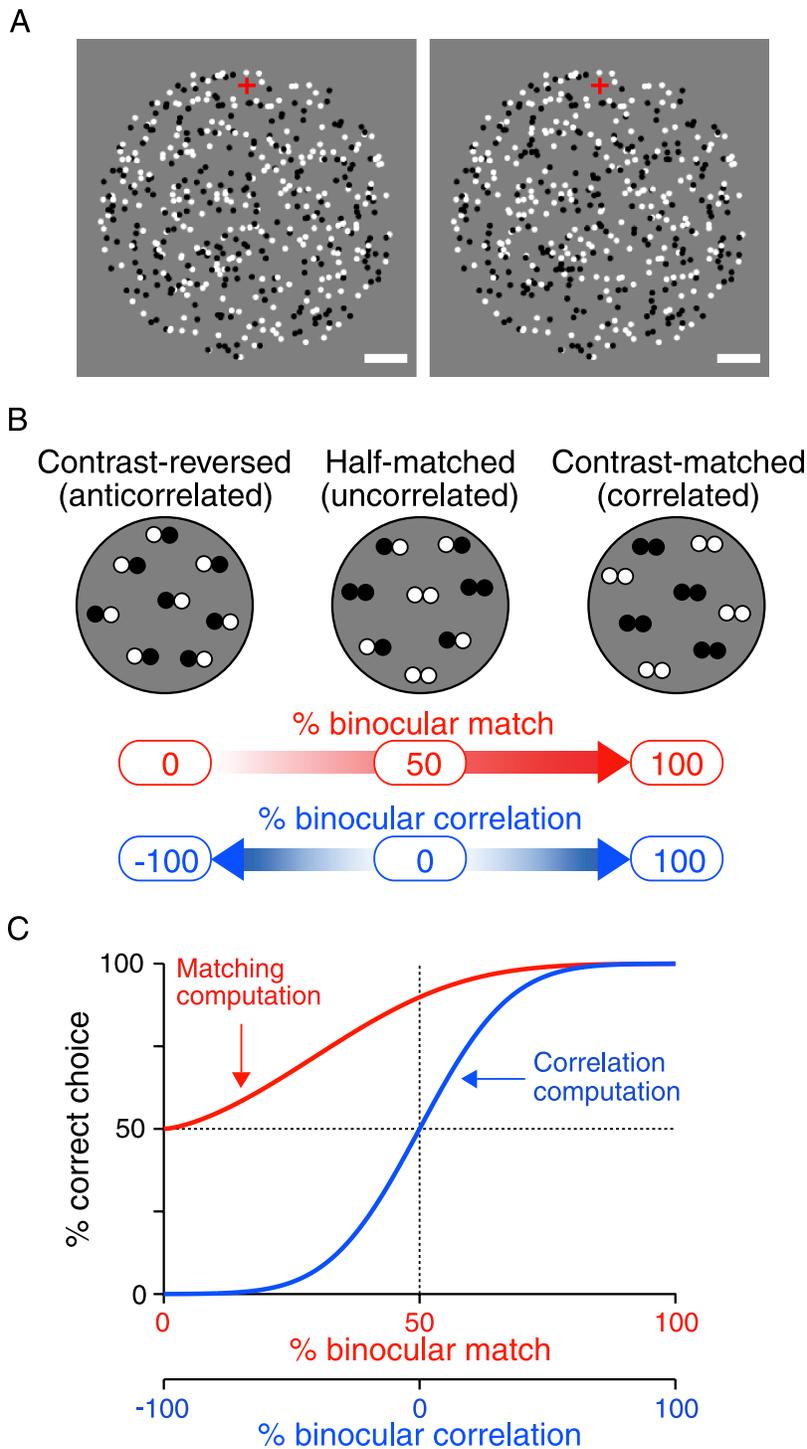
disparity tuning to anti-correlated stereograms is abolished (Janssen, Vogels, Liu, & Orban, 2003; Kumano, Tanabe, & Fujita, 2008; Tanabe et al., 2004). This suggests a matching computation in which the disparity signal increases with the percentage of matched features across the eyes. The matching computation is based on the binocularly matched features and ignores unmatched features. Overall, the stimulus correlation and matched features are -100% and 0% for anti-correlated stereograms, respectively.

How these neural computations for disparity contribute to binocular depth perception, or stereopsis, is poorly understood. To assess this, we developed random-dot stereograms (RDSs) that differentiate the stimulus strengths of the two computations. The RDSs consisted of dark and bright dots on a gray background (Figure 1A). When a proportion of dots were contrast reversed (from dark to bright or vice versa) in one eye, the stimulus strength decreased toward negative values for the correlation computation, while it decreased toward zero for the matching computation (Figure 1B). At one extreme, we used normal RDSs in which all corresponding dots had identical contrast in the two eyes. In this case, both the correlation and match levels were positive. At the other

extreme, with all dots contrast reversed, the correlation level was negative while the match level was zero. In an intermediate case, where half of the dots were contrast reversed, the correlation level was zero while the match level was positive.

The two computations give contrasting predictions of a subject’s performance in a two-alternative near/far discrimination task (Figure 1C). When required to discriminate the direction of depth (i.e., near or far) relative to the

plane of fixation, both computations produce their own decision variables by subtracting sensory responses between near and far detectors. Although cortical disparity detectors exhibit a continuous distribution in terms of their peak disparity preference and phase tuning (Cormack, Stevenson, & Schor, 1993; Cumming & DeAngelis, 2001; DeAngelis & Uka, 2003; Stevenson, Cormack, Schor, & Tyler, 1992), the activation of near or far detectors is known to have opposite effects on a decision (DeAngelis,



Cumming, & Newsome, 1998). Such opposing mechanisms successfully explain a wide range of perceptual decisions (Shadlen, Britten, Newsome, & Movshon, 1996), including stereoscopic depth discrimination (Neri, Parker, & Blakemore, 1999; Prince & Eagle, 2000; Uka & DeAngelis, 2004). According to the above explanation, the correlation computation predicts that a negative correlation results in a reversal of perceived depth because a negative correlation reverses the response balance between near and far detectors. Here, we define the correct depth of a disparity as the perceived depth of a 100% matched RDS. Using this definition, a reversal in depth perception means discriminating near/far depth correctly is below chance (<50%). The correlation computation further predicts that the percent correct (percentage of correct choices) is at chance level (50%) for a 0% correlation stimulus. This is because at 0% correlation stimulus disparity detectors based on correlation computation should lose disparity selectivity, meaning that the disparity-tuning functions are flat. Therefore,

Figure 1. Experimental rationale. (A) Example of a stereogram. A half-matched random-dot stereogram (RDS) in which half of the dots binocularly contrast-matched and the other half contrast-reversed is shown. Binocular fusion of the left and right images reveals that the RDS consists of a center disk and a surrounding annulus. In the center disk, all dots have a certain (non-zero) binocular disparity, while those in the annulus have zero disparity. Red crosses are fixation markers. During the experiments, we used red phosphors for stimuli and background while using white fixation crosses. Scale bars indicate 1° and were not shown to the subjects. (B) Graded contrast reversal differentiates binocular match and correlation. A schematic illustration of three representative stimuli is shown. Each pair of dots represents the luminance contrast of a corresponding dot in the left and right eyes. (Right) All pairs have matched contrast between the two eyes, e.g., black dots in one eye are matched with black dots in the other eye, likewise for white dots. Thus, the images from the two eyes are perfectly correlated. (Left) All pairs have reversed contrast; thus, images from the two eyes are perfectly anti-correlated. (Center) Half of the pairs have matched contrast; therefore, the images from the two eyes are uncorrelated because correlated and anti-correlated dots cancel each other. We characterized our random-dot stimuli by the percentage of pairs with matched contrast (red numbers) or the percentage of interocular correlation (blue numbers). (C) Psychometric functions predicted by the matching and correlation computations in a two-alternative near/far discrimination task. The probability of a correct choice in depth judgment is plotted against the two ways of representing the same stimuli (% binocular match and % binocular correlation). The prediction by the matching computation (red curve) is based on the stimulus representation of a given % binocular match. The prediction by the correlation computation (blue curve) is based on the stimulus representation of a given % binocular correlation. Responding correctly for 50% of the trials is denoted as chance performance.

any model solely based on correlation computation cannot explain psychophysical performances that deviate from chance level. These constraints mean that for binocular correlation, the overall percent correct should be an odd function (Figure 1C, blue curve). In contrast, the matching computation predicts that a subject will perform better than chance as long as a proportion of dots are matched between the eyes. The percent correct should decrease to chance level at the low end of the binocular match (0% match; Figure 1C, red curve). Throughout the rest of this paper, we represent our stimulus in accordance with the binocular match level.

Accordingly, human stereopsis may involve both computations. We manipulated the disparity magnitude (i.e., the absolute value of disparity) across blocks of trials in an attempt to vary the contributions of the two computations. Changing the disparity magnitude makes stereopsis transition between putative fine and coarse mechanisms (Jones, 1977; Norcia, Sutter, & Tyler 1985; Ogle, 1952; for reviews, see Bishop & Henry, 1971; Tyler, 1990; Wilcox & Allison, 2009), which may relate to the ventral versus dorsal distinction described above.

Methods

Task and testing procedure

Subjects performed a single-interval, two-alternative forced-choice near/far discrimination task. Subjects viewed a circular, bipartite dynamic RDS and judged whether the center disk was nearer or farther than the surrounding annulus (Figure 1A). In Experiment 1, we used a stimulus duration of 1.5 s such that subjects had more than enough time to make a decision. Within a block of trials, the disparity sign (uncrossed or crossed) and the percentage of binocularly contrast-matched dots (binocular match level; from 0% to 100% at 12.5% increments) were randomly ordered, whereas the disparity magnitude was kept constant. Five disparity magnitudes (0.03 , 0.06 , 0.12 , 0.24 , and 0.48°) were tested in different blocks of trials. The annulus always consisted of 100% contrast-matched dots with zero disparity except for Experiment 2. In Experiment 2, we made the surrounding annulus and the center disk have the same match level (from 0% to 100% at 12.5% increments). Two disparity magnitudes (0.03 and 0.48°) were tested. In Experiment 3, we controlled for possible vergence eye movements. The stimulus duration was 94 ms, which is too short a time period for substantial vergence eye movements to occur (Masson, Busetini, & Miles, 1997), and only 0 and 100% match levels and $\pm 0.48^\circ$ disparities were used.

Each stimulus was presented 30 times in Experiments 1 and 2 and 100 times in Experiment 3. After the subjects

performed trials for all stimulus conditions in random order, the next repetition was started. In [Experiments 1 and 2](#), a psychometric function was based on data from two blocks of 270 trials (2 disparity signs \times 9 match levels \times 15 repetitions) carried out within a single day. On each day, we chose the tested disparity magnitude randomly. In [Experiment 3](#), each data point was based on two blocks of 200 trials (2 disparity signs \times 2 match levels \times 50 repetitions).

We required subjects to quickly start and maintain their fixation during the presentation of the fixation cross. The fixation cross was shown continuously beginning 750 ms prior to the stimulus presentation until the end of the stimulus presentation. To help fixation in depth, we presented a few static random dots with zero disparity (on the screen) around the cross. After an RDS presentation, the subjects reported their depth judgment by pressing designated keys on a computer keyboard. The interstimulus interval was 2 s.

We presented the words “NEAR” or “FAR” on the screen to indicate the keyboard report made by the subjects. Subjects confirmed that they had correctly reported what they perceived with this feedback. When subjects mistakenly pressed a key that did not reflect their perceptual judgment, they were allowed to change the report until the end of each trial. However, we did not provide feedback regarding whether their response was correct.

Subjects

Four subjects (TD, TO, MH, and MT) participated in [Experiment 1](#); three subjects (TD, TO, and MT) participated in [Experiment 2](#); and six subjects (TD, TO, MH, MT, ST, and SH) participated in [Experiment 3](#). Subject TD is an author; the others were naive to the purpose of this study. All subjects had normal or corrected-to-normal vision. Before the experiments, we obtained informed written consent from all subjects.

Apparatus

RDSs were presented with a spatial resolution of 1152 \times 864 pixels at a frame rate of 85 Hz on a full-flat 17-inch CRT display (Trinitron Multiscan E230, Sony, Tokyo). The viewing distance was 57 cm. Images were anti-aliased so that visual stimuli had subpixel resolution. For dichoptic presentation of stimuli, we used liquid crystal shutter glasses (RE7-CANE, Elsa, Aachen), a graphics board supporting quad buffer stereo display (Wildcat VP 990, 3Dlabs, Milpitas, CA), and a workstation (PWS360, Dell, Round Rock, TX). The graphics library was the OpenGL Utility Toolkit. The luminance of the display background measured through the active shutter glasses was 0.95 cd/m².

Because of their short decay time, we only used red phosphors to minimize interocular crosstalk (<3% of the background).

Visual stimuli

The center disk of the concentric-bipartite RDSs had a radius of 2.5°, while the surrounding annulus had a width of 1°. To prevent any monocular features from changing systematically with disparity, we masked the positional shift of the center disk by adding or removing dots of the surrounding annulus near the disparity border while keeping the size of the center patch constant. The center of the RDSs was 3° below a fixation cross. Each eye was presented with a new dot pattern at a refresh rate of 10.6 Hz. RDSs contained an equal number of bright (1.9 cd/m²) and dark (0.01 cd/m²) dots. The dot size was 0.14 \times 0.14° with anti-aliasing. The dot density (the fraction of the non-background area to total area) was 24%. One dot occluded another dot where dots overlapped. The probability of a contrast-matched dot being occluded by a contrast-reversed dot was equal to the probability of a reversed dot being occluded by a matched dot.

Descriptive psychometric function

We fitted a psychometric function (P_d) that describes the proportion of correct choices as follows:

$$P_d(x) = 1 - (1 - \gamma) \cdot \exp\left\{-\left(\frac{x}{\alpha}\right)^\beta\right\}, \quad (1)$$

where x is the match level (percentage of contrast-matched dots), γ is the y -intercept (i.e., $P_d(0)$), α is the value of x at the psychophysical threshold such that $P_d(\alpha) = 1 - (1 - \gamma) \times 0.368$, and β is proportional to the slope at $P_d(\alpha)$. We searched for a set of parameters that maximized the likelihood of observing the data by assuming a binomial distribution. For each subject, 15 free parameters (3 parameters \times 5 psychometric functions tested at different disparity magnitudes) were required to describe the entire data set.

We measured the similarity between the observed psychometric functions and the prediction of the correlation computation (blue curve in [Figure 1C](#)). The similarity was quantified as the fractional area (F), defined as

$$F = \frac{2 \int_0^{x_c} \{0.5 - P_d(x)\} dx}{100 \int_0^{x_c} |0.5 - P_d(x)| dx}, \quad (2)$$

where x_c is the value of x satisfying $P_d(x) = 0.5$. Substituting Equation 1 gives

$$x_c = \begin{cases} \alpha \{\log(2 - 2\gamma)\}^{\frac{1}{\beta}} & \text{if } 0 \leq \gamma < 0.5 \\ 0 & \text{otherwise} \end{cases}. \quad (3)$$

The value of F measures the contribution of the odd symmetric component, centered at 50% match and 50% correct. In cases where $P_d(x)$ did not cross 0.5, x_c was set to zero. The denominator in Equation 2 is equal to the total area of the deviation from $P_d(x) = 0.5$ (Figure 3A, inset). The numerator is equal to twice the area of only the downward deviation from $P_d(x) = 0.5$. The fractional area is zero for the matching computation prediction and unity for the correlation computation prediction (Figure 1C).

Weighted average of matching and correlation signals

Near/far discrimination during a single trial

To explain the results of Experiment 1, we introduced four hierarchical stages that transformed a bivariate input, which is the disparity sign and the binocular match level, into a binary output, which is the choice of near versus far. The four stages were encoding, subtraction, weighted averaging, and binary decision (see Appendix A for details). The first stage placed the matching and correlation computations in parallel subsystems. Each subsystem consisted of a near (i.e., crossed disparity) detector and a far (i.e., uncrossed disparity) detector. The detectors had odd-symmetric tuning functions, meaning that disparities of different signs have opposite effects on the response of the detectors. This property is commonly found in the extrastriate cortices such as the middle temporal (MT) area (DeAngelis & Uka, 2003) and V2 (Tanabe & Cumming, 2008). For simplicity, step functions were used for the disparity-tuning functions. The detectors changed their responses only when the disparity switched signs. Alternatively, we used a set of Gabor functions with various horizontal offsets instead of a single step function to make the disparity detectors more physiologically relevant.

Monotonic functions of the match level dictated the detectors' tuning amplitude. In the correlation subsystem, the function was linear, while in the matching computation it was sigmoid, as these functions best described the data (see Appendix A for details). We assumed independent Gaussian noise in the responses of the detectors, with the noise variance proportional to the mean (Dean, 1981). The later stages involved the decision-making processes. The second stage subtracted the output of the near detector from that of the far detector for each subsystem. The third stage calculated a decision variable by using a weighted average of the signals across the subsystems.

The fourth stage made the binary decision according to the sign of the decision variable.

Psychometric functions and fitting procedures

We derived the psychometric function P_w from the weighted average of the two computations (see Appendix A for derivation) such that

$$P_w(x) = \frac{1}{2} \left\{ 1 + \text{Erf} \left(\frac{a \{ w f_1(x) + (1-w) f_2(x) \}}{\sqrt{w^2 + (1-w)^2}} \right) \right\}, \quad (4)$$

where Erf denotes the error function, a is the response amplitude of the encoding detectors, w is the relative weight of the correlation computation over the matching computation, and $f_1(x)$ and $f_2(x)$ describe the dependency of the detectors' responses on the match level x for the correlation computation and for the matching computation, respectively. $f_1(x)$ linearly transforms the match level x into a value ranging from -1 to 1 (Equation A2). $f_2(x)$, which has two parameters (u and l), is a sigmoidal function that transforms x into a value ranging from 0 to 1 (Equation A3). We fitted this function to the data obtained in the five experiments using different disparity magnitudes while solving for a , u , and l over all five experiments and w for each experiment, resulting in a total of eight free parameters (a , u , l , and five weights, w_1 , w_2 , w_3 , w_4 , and w_5) for each subject. As in the fit of the descriptive psychometric function (Equation 1), we used the maximum likelihood estimation and assumed a binomial distribution for the observed number of correct choices. All data analyses were done with MATLAB (Mathworks, Natick, MA).

Results

Experiment 1: Fine and coarse near/far discriminations follow different psychometric functions

When the subjects discriminated the direction of fine depth at a disparity magnitude of 0.03° , the psychometric functions agreed with the prediction of the matching computation (solid curves in Figure 2). As the binocular match level was lowered from 100% to 0%, the percent correct decreased from perfect to chance level. When subjects discriminated the direction of coarse depth at a disparity magnitude of 0.48° , the psychometric functions were mixtures of matching and correlation computation predictions (dashed curves in Figure 2). The percent

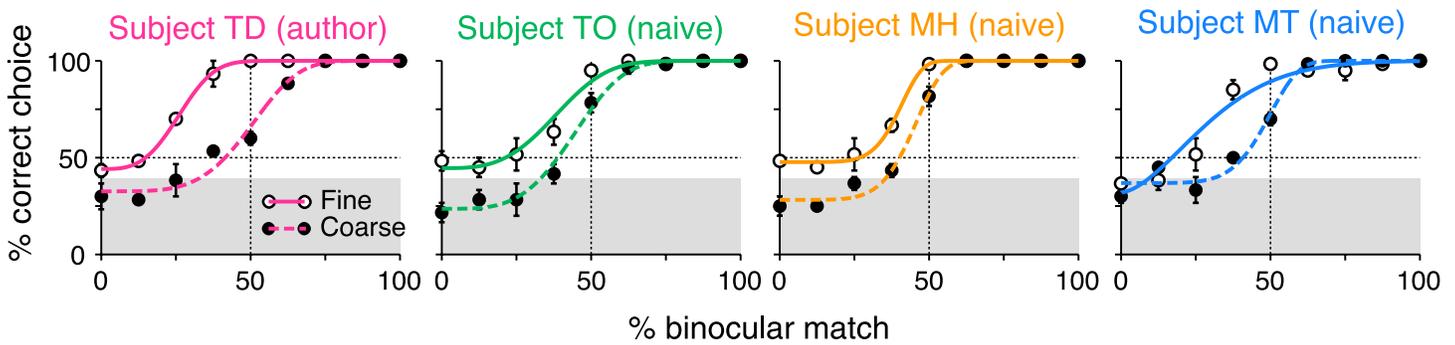


Figure 2. Psychometric functions in fine and coarse near/far discriminations have different shapes (Experiment 1). The percentages of correct choices from four subjects (one author and three naives) in fine (0.03°) and coarse (0.48°) near/far discrimination tasks are plotted as a function of the % binocular match. Continuous and dashed lines represent the best-fitted descriptive psychometric functions (Equation 1) using fine and coarse discrimination data, respectively. The shaded gray area indicates the region where the percentage of correct choices is significantly below chance performance ($p < 0.05$, binomial test). Each data point shows the mean calculated from 60 choices. Error bars indicate standard errors of the means (SEMs) across two blocks of trials.

correct for coarse depth was lower than that for fine depth at low match levels ($\leq 50\%$) and fell below chance level with some of the lowest match levels. We confirmed that the significantly low performance was not an artifact of vergence eye movements (see results of Experiment 3). Data at intermediate disparity magnitudes are available in Supplementary Figure S1.

We calculated the fractional area (Equation 2) to quantify how well the observed psychometric function follows the prediction of the correlation computation (Figure 3A). The fractional area was small at the smallest disparity (0.03° ; mean \pm standard deviation (SD), 0.06 ± 0.037 , $n = 4$), but much larger at the largest disparity (0.48° ; mean \pm SD, 0.42 ± 0.074 , $n = 4$). The fractional area increased with the disparity magnitude (Figure 3B; regression slope of 0.32 for data pooled across all subjects, $p = 2.8 \times 10^{-6}$, H_0 : linear-regression slope against the common log of the disparity magnitude is 0). In two subjects (TO and MT), the fractional area increased gradually, while in the other two (TD and MH) it abruptly increased at 0.2° . As the disparity magnitude increased, the psychometric function changed from following the prediction of the matching computation to partially following the prediction of the correlation computation.

Care is needed when interpreting the observed shift of the psychometric functions because the disparity magnitude might affect task difficulty. For example, the task is particularly difficult for subjects if disparities approach the stereoacuity threshold. As the task becomes easier with larger disparity, deviation from chance should increase, but there should be no shift in x_c (the match level at chance performance). In contrast with these predictions, x_c increased as the task became easier (Figure 4A; regression slope of 20.1, $p = 1.2 \times 10^{-3}$), although the increase was not monotonic in two subjects (TD and MH). The high performance at 50% match decreased toward chance as the task became easier (Figure 4B; regression slope of -13.7 , $p = 0.0051$). We found only one tendency that

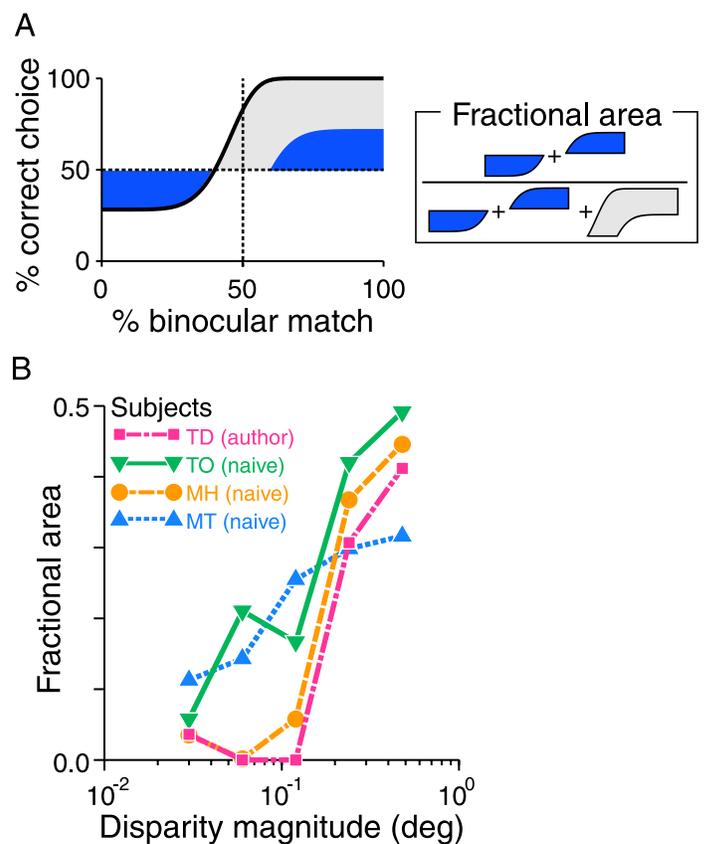


Figure 3. Relationship between the psychometric function and the disparity magnitude (Experiment 1). (A) The shape of the psychometric function was assessed by decomposing the deviation from chance performance into areas representing the odd-symmetric component (blue area) and everything else (gray area). The fractional area is the area of the odd-symmetric component divided by the area of the net deviation (Equation 2). (B) The fractional area is plotted against the log disparity magnitude for four subjects. The fractional area increased with the disparity magnitude.

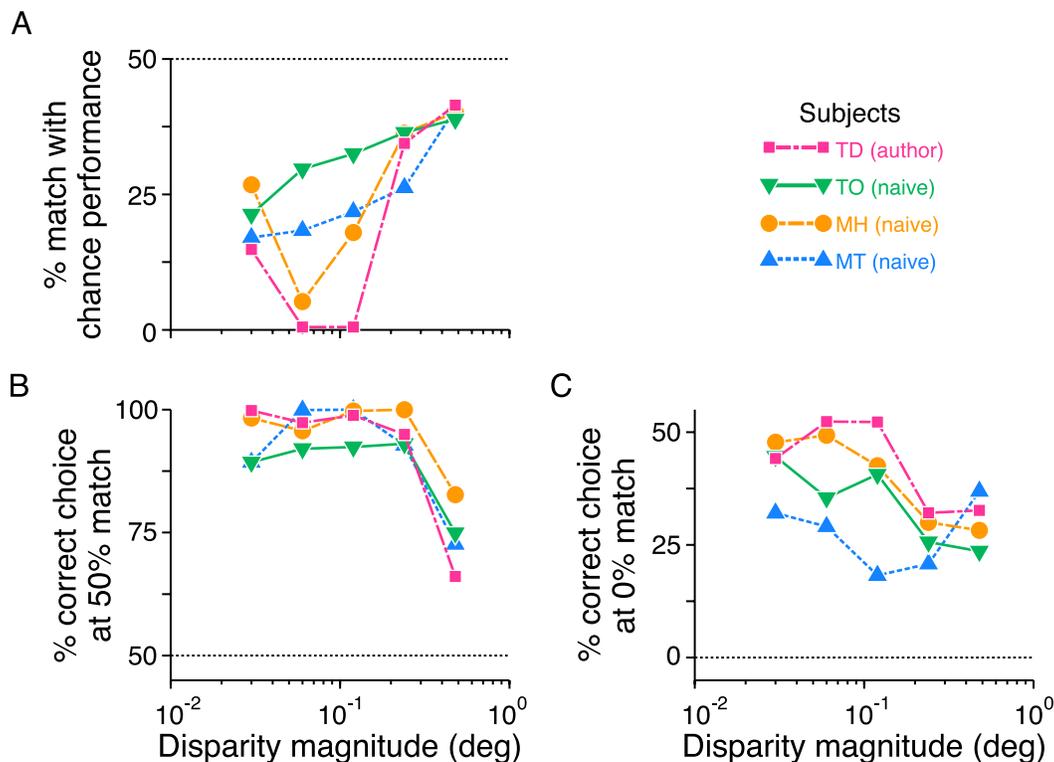


Figure 4. Further quantitative assessments of the psychometric function (Experiment 1). Three quantities are plotted against the log disparity magnitude for four subjects. (A) Binocular match levels with chance performance increased with disparity magnitude. (B) The percentage of correct choices with 50% match decreased. (C) The percentage of correct choices with 0% match decreased. We used the fitted psychometric functions to estimate these quantities. The horizontal dotted lines indicate predictions by the correlation computation.

agreed with the expected results if task difficulty was the cause; the performance at 0% match deviated away from chance as the task became easier (Figure 4C; regression slope of -12.6 , $p = 0.016$). Nevertheless, in general, task difficulty does not explain the observed shift of the psychometric functions.

Experiment 2: Effects of the binocular match level of the surrounding annulus

Studies on stereo psychophysics disagree as to whether stereoscopic depth is compatible with matching computation (Cogan, Lomakin, & Rossi, 1993; Cumming, Shapiro, & Parker, 1998; Julesz, 1971; Read & Eagle, 2000) or with correlation computation (Cormack, Stevenson, & Schor, 1991; Neri et al., 1999; Rogers & Anstis, 1975; Tanabe, Yasuoka, & Fujita, 2008). Key evidence supporting the correlation computation is reversed depth with a contrast-reversed stereogram. In a previous study, our group has shown that reversed depth with contrast-reversed RDSs is abolished when a crisp reference is not available (Tanabe et al., 2008). The psychometric functions in Experiment 1 might similarly depend on the availability of a 100% match reference. To examine this possibility, we made the surrounding annulus and the center disk in Experiment 2 have the same match level.

We tested 0.03° (fine) and 0.48° (coarse) disparity magnitudes. Other experimental conditions were the same as Experiment 1. For all three subjects, coarse near/far discrimination was strongly affected by the match level of

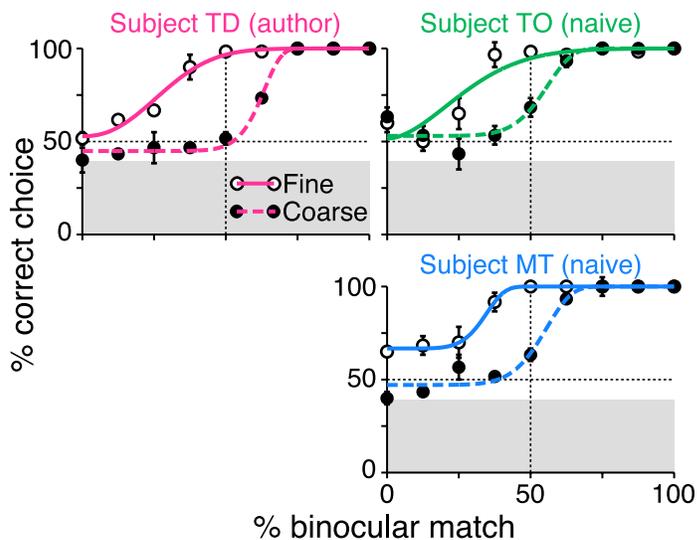


Figure 5. Fine and coarse near/far discriminations when the surrounding annulus has the same binocular match level as the center disk (Experiment 2). The conventions are the same as in Figure 2. Each data point is based on 60 choices.

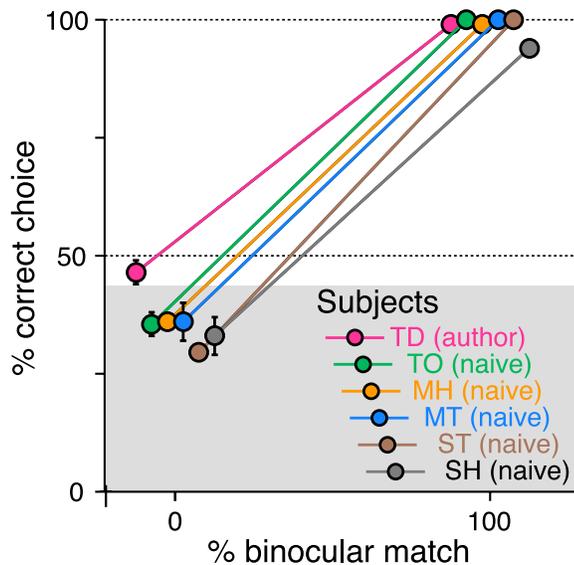


Figure 6. Coarse near/far discrimination with short-duration (94 ms) stimulation (Experiment 3). Contrast-matched RDSs (100% binocular match) and contrast-reversed RDSs (0% binocular match) with either 0.48° or -0.48° disparity were presented 100 times in random order in two blocks of trials. The percentage of correct choices among 200 choices (choices were pooled across disparity signs) is plotted for contrast-reversed RDSs (0% match) and contrast-matched RDSs (100% match). Error bars represent SEMs across two blocks of trials. In the shaded area, the data points are significantly lower than chance performance ($p < 0.05$, binomial test).

the surround such that reversed depth perception was abolished (compare Figure 5 with Figure 2). In the range where the match level was between 0% and 50% (from -100% to 0% correlation level), the psychometric functions were almost flat near chance performance. The results support the hypothesis by Tanabe et al. (2008) that reversed depth perception requires a reference stimulus in which the majority of elements are contrast-matched between the two eyes.

Experiment 3: Effects of vergence eye movements

Vergence eye movements are another potential artifact (Masson et al., 1997; Stevenson, Cormack, & Schor, 1994; Takemura et al., 2001). A reversal of vergence might cause the reversal of depth perception, even without correlation computation. When convergence occurs in response to a far disparity with a 0% match stimulus, the disparity of the surrounding annulus becomes uncrossed. The decision circuit might then exploit the signal of the surround to infer the ambiguous depth of the center. The

inferred depth is presumably near because a near choice is associated with an uncrossed surround when the stimulus is a 100% match.

We controlled for vergence eye movements by presenting visual stimuli briefly for only 94 ms. This duration is too short a time period for substantial vergence eye movements to occur (Masson et al., 1997). Contrast-matched RDSs (100% match level) and contrast-reversed RDSs (0% match level) with either 0.48° or -0.48° disparity were presented. The percentage of correct choices in five of the six subjects with contrast-reversed RDSs was significantly below 50% (Figure 6; subject TD, $p = 0.18$; subject TO, $p = 2.5 \times 10^{-5}$; subject MH, $p = 4.6 \times 10^{-5}$; subject MT, $p = 4.6 \times 10^{-5}$; subject ST, $p = 3.2 \times 10^{-9}$; subject SH, $p = 8.7 \times 10^{-7}$; binomial test), indicating that the five subjects perceived reversed depth for contrast-reversed RDSs. The performance of the remaining subject (TD) with contrast-reversed RDSs was also numerically lower than 50%. The probability of observing such bias across six subjects by chance is 0.016 ($1/2^6$). Therefore, reversed depth in contrast-reversed RDSs was unlikely to be caused by vergence eye movements, suggesting that correlation computation contributes to stereopsis without the execution of vergence eye movements.

Weighted average of matching and correlation signals explains the changes of the psychometric functions

A single factor, the relative weight of matching and correlation signals, was found to explain how the psychophysical performance depends on the disparity magnitude (i.e., the results of Experiment 1) by deriving the psychometric function from the weighted average of the matching and correlation signals (Figure 7A; Equations 4, A2, and A3). The relative weight was the only parameter that varied with disparity magnitude. We found that the fitted function successfully captured how the data depends on the disparity magnitude (Figure 7B). The fit of the function was as good as the individual fits of the descriptive functions shown in Figure 2 (Figure 7C; normalized log likelihood, mean \pm SD, $96 \pm 2.6\%$, $n = 4$). The weight of the correlation computation relative to that of the matching computation increased with the disparity magnitude (Figure 7D; regression slope of 0.22, $p = 1.0 \times 10^{-4}$) in a manner similar to the fractional area increase ($r = 0.88$, $p = 3.2 \times 10^{-7}$). We could replicate the above results even when using Gabor function disparity tunings with various preferred disparities (normalized log likelihood, $95 \pm 3.1\%$; regression slope of relative weight, 0.13, $p = 6.3 \times 10^{-4}$; correlation of relative weight between the two fits, $r = 0.85$, $p = 2.2 \times 10^{-6}$). The incorporation of Gabor functions makes the analysis more

physiologically relevant. Overall, most changes in the psychometric function were explained by changing the relative weight of the two computations.

Discussion

The subjects' near/far discrimination at fine disparity agreed with the prediction made using matching computation. However, at large disparity, the subjects' performance deviated from the prediction. The data were consistent with the decision process being described by a weighted average of the matching and correlation computations. The relative weight of the correlation computation increased with the disparity magnitude. Therefore, we conclude that the matching and correlation computations support stereoscopic depth perception with varying contributions depending on the disparity magnitude.

Contributions of matching and correlation computations to stereopsis

This study provides an answer to why some stereo depth perception studies have found evidence for correlation computation while others have not. We confirmed our previous hypothesis that for stereo depth perception to rely on the correlation computation, certain conditions must be met (Tanabe et al., 2008). One crucial condition is that the reference plane has binocularly matched dots (Figure 5; Tanabe et al., 2008). An earlier study has proposed other conditions like low luminance contrast images (Cormack et al., 1991).

We found that stereo processing transitions from a regime in which the matching computation dominates perceptual judgment to a regime in which the correlation computation contributes as the disparity magnitude increases (Figures 2 and 3). The psychometric functions underwent systematic shifts that were best described by the weighted average of the two computations (Figure 7). Changes in task difficulty were ruled out (Figures 4A and 4B). Of the various parameters in the weighted-average model, the relative weight of the two computations proved able to describe the systematic shifts in the data (Figure 8D). In contrast, the amplitude of the detectors did not shift x_c , the point of intersection with chance performance (Figure 8A), while varying the parameters u and l (Equation A3) did not shift $P_w(0)$, the y -intercept (Figures 8B and 8C).

However, the weighted average of the two computations does not explain why the reference plane must be binocularly matched for the correlation computation to yield reversed depth. The stereoscopic system might have a gating process that passes reversed depth signals from the correlation computation only when the reference plane

is binocularly matched. This process can be implemented by a thresholding operation where the signals can exceed the threshold when a matched reference adds a positive offset to the signals but not when an unmatched reference adds only a small offset. Because disparity encoding is independent between the stimulus center and surround in the primary visual cortex, such signal combination across the stimulus center and surrounding reference is likely to occur in extrastriate visual areas (Cumming & Parker, 1999).

Can the correlation computation explain fine discrimination data?

The data presented in this paper cannot be explained solely by the correlation computation. Population codes based on the correlation computation often encounter conflicting signals between spatial frequency channels, preferred disparities, and tuning symmetries. Although these conflicting signals can explain chance performance with 0% match stimuli, they cannot explain high performance with 50% match stimuli. To correctly interpret the conflicting signals, it has been proposed that only signals that do not conflict are passed to later visual stages (Read, 2002; Read & Eagle, 2000). No matter how the conflict is resolved, their model can only perform at chance with 50% match because all detectors are signaling at their baseline, which strays significantly from the experimental data. This holds for any correlation-based detectors that have different tuning widths, preferred disparities, or tuning symmetries. Therefore, correlation computation alone is insufficient for explaining our psychophysical data.

Neural substrates for fine and coarse stereopsis

Early studies of disparity signals in V1 asserted a relationship between disparity-tuning categories and fine/coarse stereopsis (Poggio & Fischer, 1977; Poggio, Gonzalez, & Krause, 1988). This assertion is based on very specific assumptions: detectors tuned to small disparities should have a particular property related to and send the signals to a particular circuit for fine stereopsis. Likewise, detectors tuned to large disparities should have a similar effect on coarse stereopsis. However, no evidence exists supporting these assumptions.

It is likely that not just one but two or more cortical areas feed disparity signals for fine and coarse stereopsis (Neri, 2005; Orban et al., 2006; Parker, 2007). Our results suggest that stereopsis combines disparity signals from the two distinct computations. There is growing evidence that the cortical areas in the ventral pathway subserve the matching computation (Janssen et al., 2003; Kumano et al.,

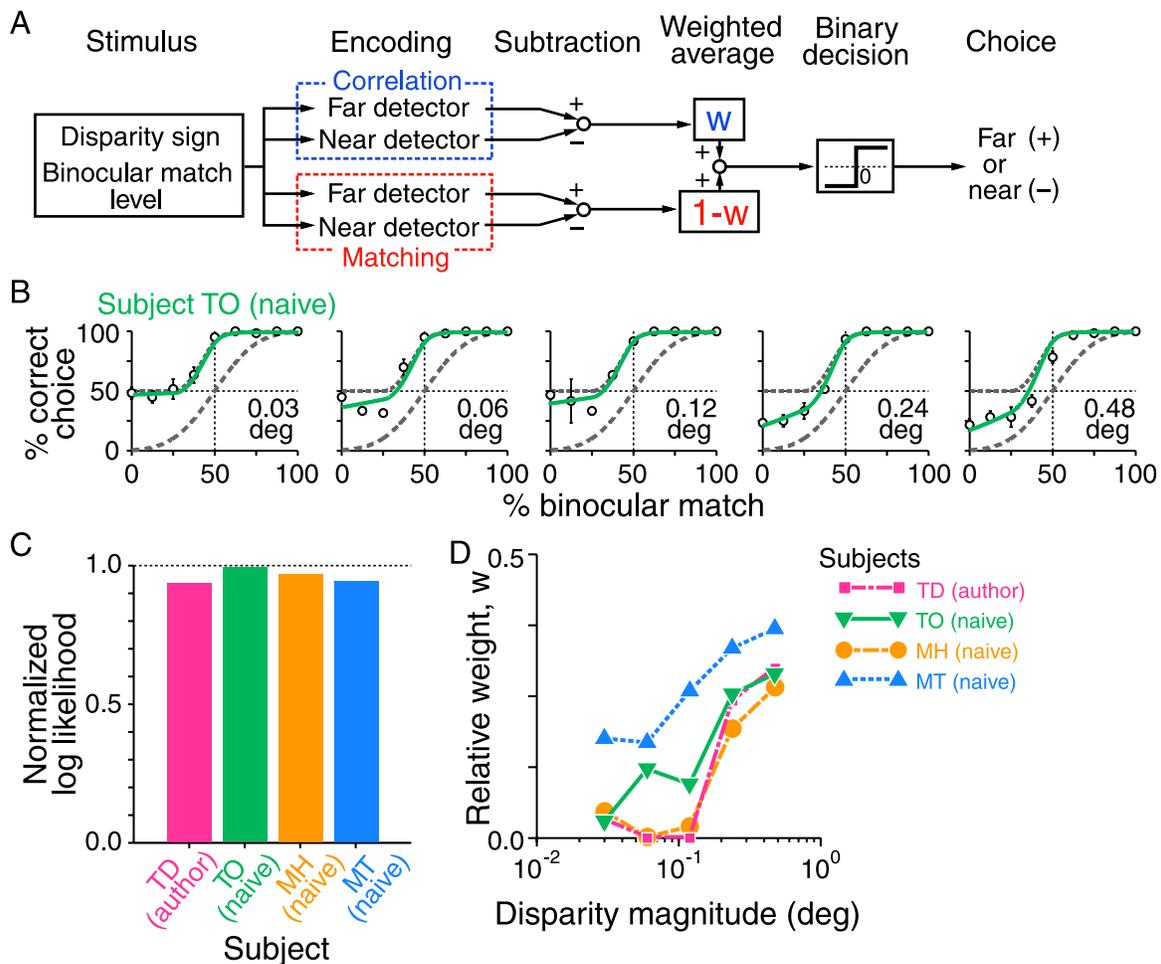


Figure 7. Weighted average of matching and correlation signals (an explanation for the results of Experiment 1). (A) A schematic diagram is shown. A bivariate input is transformed into a binary choice. The transformation involves four stages. The parameter w controls the relative contribution of the correlation computation for depth judgment. (B) The best-fitted functions with only w fitted independently across five disparity magnitudes (solid curves, Equations 4, A2, and A3). The tested disparity magnitude is shown in each panel. The two dashed curves above and below the solid curves are the hypothetical psychometric functions for pure matching computation ($w = 0$) and pure correlation computation ($w = 1$), respectively. Each data point is based on 60 choices. Error bars indicate SEMs across two blocks of trials. (C) The log likelihood of the fits was plotted on a scale that ranges from zero (i.e., the log likelihood of a random choice) to one (i.e., the log likelihood of descriptive psychometric functions; also see Figure 2 and Equation 1). (D) The relative weight of the correlation computation, w , was plotted against the log disparity magnitude for each subject.

2008; Tanabe et al., 2004), while the areas in the dorsal pathway subserve the correlation computation (Krug et al., 2004; Takemura et al., 2001; see Preston, Li, Kourtzi, & Welchman, 2008 for an opposing perspective). Therefore, fine depth perception may rely mainly on the ventral pathway, while coarse depth perception may rely on both. This view agrees with physiological evidence for the perceptual roles of the two pathways. In the dorsal pathway, microstimulation applied to area MT biased behavioral judgment of coarse but not fine depth (Uka & DeAngelis, 2006), while in the ventral pathway, inferior temporal (IT) neurons showed trial-by-trial response variation correlated with fine depth judgment (Uka, Tanabe, Watanabe, & Fujita, 2005). Our view also agrees

with a neurological study with a patient (DF) whose lateral occipital cortex and ventral pathway functions are damaged (Read, Phillipson, Serrano-Pedraza, Milner, & Parker, 2010). This patient performed poorer than control subjects particularly when making psychophysical judgment based on fine depth information.

Mechanisms mediating match-based signals

Area V1 generates the correlation-based signals in the form of disparity energy (Cumming & Parker, 1997; Ohzawa, DeAngelis, & Freeman, 1990). The pathway leading to the ventral areas transforms the disparity energy

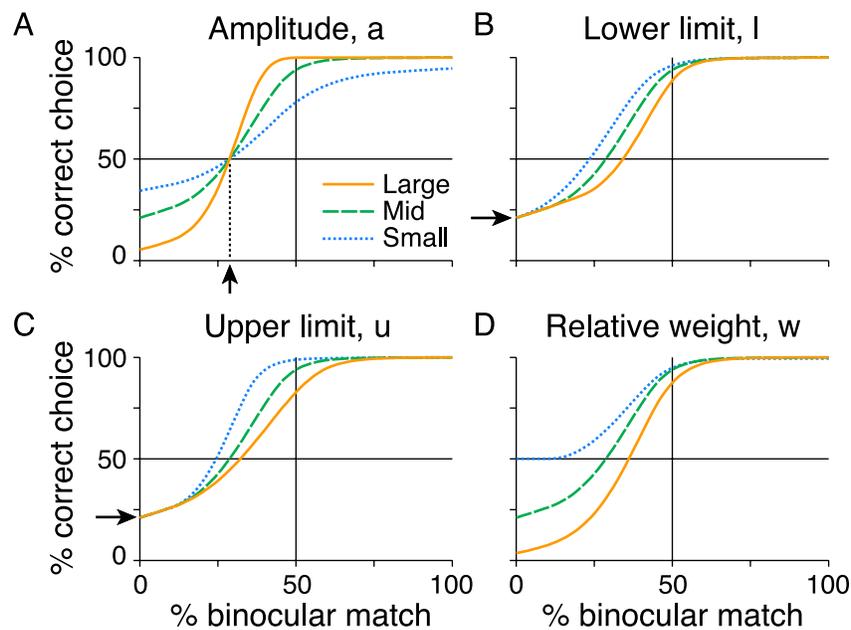


Figure 8. The influence of parameters on the weighted-average psychometric function. (A) Psychometric functions (Equations 4, A2, and A3) are shown as the amplitude parameter a is gradually increased from small (blue), mid (green), to large (orange). The black arrow indicates an anchored point. (B) Psychometric functions for when the lower limit (l) of a non-linearity dynamic range in the matching computation gradually increases. (C) Psychometric functions for when the upper limit (u) of a non-linearity dynamic range in the matching computation gradually increases. (D) Psychometric functions for when the relative weight (w) of the correlation computation is increased.

signals to match-based signals (Janssen et al., 2003; Kumano et al., 2008; Tanabe et al., 2004). Previous studies have postulated that the signal transformation can be achieved by an additional non-linearity (Lippert & Wagner, 2001). Their model used even-symmetric tuning functions, i.e., the classic tuned-excitatory functions. To extend this model to other tuning functions, the model output needs to be further combined in a specialized manner (Haefner & Cumming, 2008; Tanabe & Cumming, 2008).

Although the two models in the previous paragraph were originally intended for 100% and 0% matches only, our simulation indicates that they fulfill the matching computation at all % match levels (data not shown). This suggests that the detectors in our weighted-average model can be replaced entirely with signals from disparity energy mechanisms. Our weighted-average model can be extended to receive stereo images rather than the original bivariate values as input, making it directly testable in physiological experiments. It would be interesting to see whether the gain of the disparity signals is counter-balanced between the dorsal and ventral pathways, as our weighted-average model predicts.

Conclusion

Two distinct computations feed the disparity signals for stereoscopic depth perception. One computes disparity

based on binocularly matched patterns, while the other computes the cross-correlation of binocular images. These signals are used differently depending on the disparity magnitude. For small disparities, signals from matched patterns dominate. For large disparities, the contribution of the correlation computation increases. These findings shed light on an apparent contradiction in the literature in which some studies suggest that stereoscopic depth perception relies on a matching computation (Cogan et al., 1993; Cumming et al., 1998; Julesz, 1971; Read & Eagle, 2000), while others support a correlation computation (Cormack et al., 1991; Neri et al., 1999; Rogers & Anstis, 1975; Tanabe et al., 2008).

Appendix A

Derivation of the psychometric function from the weighted average of the two computations (Equation 4)

The first encoding stage consists of two parallel subsystems: one performing the matching computation; the other performing the correlation computation. These two subsystems encode the disparity sign s (-1 for “near” disparity and $+1$ for “far” disparity) with different dependencies on the % binocular match level x . The disparity magnitude is not given as an input. However,

one parameter, the weight w , can be assigned different values depending on the disparity magnitude.

Each subsystem contains two sensory detectors: one responds to “near” disparity and is inhibited by “far” disparity, the other responds to “far” disparity and is inhibited by “near” disparity (odd-symmetric disparity tuning). The responses (R_{ij}) of these detectors are defined as follows:

$$R_{ij}(x, s) = c_j a s f_i(x) + b + \varepsilon_{ij}, \quad (\text{A1})$$

where the subscripts i and j are binary values identifying the subsystem and the detector, respectively. Throughout this section, $i = 1$ denotes the subsystem of the correlation computation and $i = 2$ denotes the subsystem of the matching computation, $j = 1$ denotes the near detector and $j = 2$ denotes the far detector. The coefficient c_j is defined as $(c_1, c_2) = (-1, +1)$. The correspondence between the coefficient c_j and disparity sign s determines the sign of the first term. Parameter a is the response amplitude. The function $f_i(x)$ maps the match level x to a value between -1 and 1 ($i = 1$, the correlation computation) or a value between 0 and 1 ($i = 2$, the matching computation). Parameter b is the response baseline. ε_{ij} represents noise and is the only random variable; all other parameters are deterministic. We assume that the response amplitude, a , and baseline, b , are the same between the two subsystems.

The mapping function of the correlation subsystem $f_1(x)$ and that of the matching subsystem $f_2(x)$ are

$$f_1(x) = \frac{x}{50} - 1, \quad (\text{A2})$$

$$f_2(x) = \begin{cases} 0 & \text{if } x < l \\ \frac{2}{(u-l)^2}(x-l)^2 & \text{if } l \leq x < 0.5(u+l) \\ -\frac{2}{(u-l)^2}(x-u)^2 + 1 & \text{if } 0.5(u+l) \leq x < u \\ 1 & \text{otherwise} \end{cases}, \quad (\text{A3})$$

where the function $f_1(x)$ is a linear map, and $f_2(x)$ is a sigmoidal function that consists of four parts: zero, expansive non-linearity, compressive non-linearity, and one. Expansive and compressive non-linearities are odd symmetric with each other. The parameters u and l determine the upper and lower limits of the dynamic range for $f_2(x)$, respectively. u was set to be larger than l .

The distribution of ε_{ij} is defined as Gaussian noise centered at zero. The noise variance is assumed to be

proportional to the mean responses of the detectors with a scaling factor φ (Dean, 1981). Thus,

$$E[\varepsilon_{ij}] = 0, \quad (\text{A4})$$

$$V[\varepsilon_{ij}] = \varphi E[R_{ij}] = \varphi \{c_j a s f_i(x) + b\}, \quad (\text{A5})$$

where E and V indicate the expectation and variance of the random variables, respectively. We assume that φ is the same between the correlation and matching subsystems.

The second stage subtracts the responses from the near ($j = 1$) and far ($j = 2$) detectors. The output of the second stage $S_i(x, s)$ is

$$\begin{aligned} S_i(x, s) &= R_{i2}(x, s) - R_{i1}(x, s) \\ &= c_2 a s f_i(x) + \varepsilon_{i2} - c_1 a s f_i(x) - \varepsilon_{i1} \\ &= 2 a s f_i(x) + (\varepsilon_{i2} - \varepsilon_{i1}) \end{aligned} \quad (\text{A6})$$

The noise distribution of the second-stage output also becomes Gaussian noise centered at zero. Its variance becomes the sum of the response variances of the near and far detectors:

$$\begin{aligned} E[\varepsilon_{i2} - \varepsilon_{i1}] &= E[\varepsilon_{i2}] - E[\varepsilon_{i1}] \\ &= 0 \end{aligned}, \quad (\text{A7})$$

$$\begin{aligned} V[\varepsilon_{i2} - \varepsilon_{i1}] &= V[\varepsilon_{i2}] + V[\varepsilon_{i1}] \\ &= \varphi \{c_2 a s f_i(x) + b + c_1 a s f_i(x) + b\} \\ &= 2\varphi b \end{aligned} \quad (\text{A8})$$

The third stage calculates the weighted average of the outputs from the matching ($i = 2$) and correlation ($i = 1$) subsystems. This is the decision variable $D(x, s)$:

$$\begin{aligned} D(x, s) &= w S_1(x, s) + (1 - w) S_2(x, s) \\ &= w \{2 a s f_1(x) + (\varepsilon_{12} - \varepsilon_{11})\} \\ &\quad + (1 - w) \{2 a s f_2(x) + (\varepsilon_{22} - \varepsilon_{21})\}, \\ &= 2 a s \{w f_1(x) + (1 - w) f_2(x)\} + w(\varepsilon_{12} - \varepsilon_{11}) \\ &\quad + (1 - w)(\varepsilon_{22} - \varepsilon_{21}) \end{aligned} \quad (\text{A9})$$

where the weight parameter w controls the relative contribution of the correlation and matching subsystems to the decision variable. Again, the noise distribution of the decision variable becomes Gaussian noise centered at zero. The noise variance is a weighted summation of the variances of the second-stage outputs from the matching and correlation subsystems. The weights are squared when the variance of a summed distribution is calculated. Thus, the distribution of the decision variable $D(x, s)$ becomes

$$E[D(x, s)] = 2as\{wf_1(x) + (1-w)f_2(x)\}, \quad (\text{A10})$$

$$\begin{aligned} V[D(x, s)] &= V[w(\varepsilon_{12} - \varepsilon_{11}) + (1-w)(\varepsilon_{22} - \varepsilon_{21})] \\ &= w^2V[\varepsilon_{12} - \varepsilon_{11}] + (1-w)^2V[\varepsilon_{22} - \varepsilon_{21}] \\ &= 2\varphi b\{w^2 + (1-w)^2\} \end{aligned} \quad (\text{A11})$$

The final stage is a binary decision, which simply passes the decision variable $D(x, s)$ through a step function:

$$T = \begin{cases} +1 & \text{if } D > 0 \\ -1 & \text{otherwise} \end{cases}, \quad (\text{A12})$$

where $T = +1$ denotes a far choice, and $T = -1$ indicates a near choice. The output is correct if choice T and the stimulus disparity sign s have the same sign.

Because the noise scaling (φ) and baseline (b) parameters are redundant (Equation A11), we set $\varphi = 1$. Then, the amplitude (a) and baseline (b) parameters are also redundant because they determine the Gaussian center and width, respectively; increasing (or decreasing) the Gaussian center and decreasing (or increasing) the Gaussian width have the same effects on the psychometric function. Therefore, we fix $b = 1$. Thus, the variance of the decision variable $D(x, s)$ can be rewritten as

$$V[D(x, s)] = 2\{w^2 + (1-w)^2\}. \quad (\text{A13})$$

From $D(x, s)$, we can calculate the probability of a correct choice $P_w(x)$. We redefine the expectation of the decision variable for a “far” disparity stimulus $E[D(x, +1)]$ as $\mu(x)$. Hence, the expectation for a “near” disparity $E[D(x, -1)]$ is rewritten as $-\mu(x)$. The variance $V[D(x, s)]$

is replaced with σ^2 . Following these changes, the probability of a correct choice can be expressed as

$$\begin{aligned} P_w(x) &= P(T = +1, s = +1) + P(T = -1, s = -1) \\ &= P(T = +1|s = +1)P(s = +1) \\ &\quad + P(T = -1|s = -1)P(s = -1) \\ &= \frac{1}{2}\{P(T = +1|s = +1) + P(T = -1|s = -1)\} \\ &= \frac{1}{2}\{P(D > 0|s = +1) + P(D \leq 0|s = -1)\} \\ &= \frac{1}{2}\left\{\int_0^{\infty} N(z|\mu(x), \sigma^2)dz + \int_{-\infty}^0 N(z|-\mu(x), \sigma^2)dz\right\} \\ &= \int_0^{\infty} N(z|\mu(x), \sigma^2)dz, \end{aligned} \quad (\text{A14})$$

where $N(z|\mu(x), \sigma^2)$ denotes a normal distribution with mean $\mu(x)$ and variance σ^2 . P indicates probability. If we transform z into t such that

$$t = \frac{z - \mu(x)}{\sqrt{2\sigma^2}}, \quad (\text{A15})$$

the psychometric function $P_w(x)$ can be rewritten as

$$\begin{aligned} P_w(x) &= \int_{-\frac{\mu(x)}{\sqrt{2\sigma^2}}}^{\infty} N\left(t|0, \frac{1}{2}\right)dt \\ &= \frac{1}{2} + \int_0^{\frac{\mu(x)}{\sqrt{2\sigma^2}}} N\left(t|0, \frac{1}{2}\right)dt \\ &= \frac{1}{2}\left\{1 + \text{Erf}\left(\frac{\mu(x)}{\sqrt{2\sigma^2}}\right)\right\}, \end{aligned} \quad (\text{A16})$$

where Erf indicates the error function. Substituting $s = 1$, $\varphi = 1$, and $b = 1$ into Equations A10 and A11 gives the final solution of the psychometric function:

$$P_w(x) = \frac{1}{2}\left\{1 + \text{Erf}\left(\frac{a\{wf_1(x) + (1-w)f_2(x)\}}{\sqrt{w^2 + (1-w)^2}}\right)\right\}. \quad (\text{A17})$$

The psychometric function has four parameters: amplitude (a), relative weight (w), and two non-linearity parameters u and l in $f_2(x)$ (see Equation A3). In the fitting procedure,

a , u , and l were kept the same in the five experiments using different disparity magnitudes. The weight parameter w was varied, however. A total of eight free parameters were used to explain the entire data set for each subject (a , u , l and w_1, w_2, \dots, w_5). The parameter a was constrained to be positive. The weight parameters (w_1, w_2, \dots, w_5) were constrained between 0 and 1. The parameters u and l were constrained between 0 and 100 with $u > l$.

Before defining $f_2(x)$ as a non-linear function, we began by defining $f_2(x)$ as a linear function that transforms the match level x into a value ranging from 0 to 1. This definition makes $f_2(x)$ more comparable to $f_1(x)$. However, this version could not explain the psychometric data, especially for fine near/far discrimination (open circles in [Figure 2](#)). The linear $f_2(x)$ yields psychometric functions steeply rising at around 0% match level, although the percentage of correct choices by the subjects gradually increased. The percentages were almost flat around chance performance between 0% and 25% match levels for subjects TO and MH. Thus, $f_2(x)$ was extended to a non-linear function. In contrast, $f_1(x)$ was kept linear because this minimal version could explain the data well and is consistent with the disparity energy model.

To make the disparity detectors more physiologically relevant, we next applied Gabor functions to the detectors' disparity tuning. We used 10 detectors for near-preferring and far-preferring populations, respectively. The k th detector of near ($j = 1$) and far ($j = 2$) population had the following tuning function, G_k^j , for disparity (d):

$$G_k^j(d) = \exp\left\{-\frac{(d - o_k^j)^2}{2\xi^2}\right\} \cdot \cos\left\{2\pi q(d - o_k^j) + \theta_j\right\}, \quad (\text{A18})$$

where o_k^j and ξ indicate the center position and size of the Gaussian envelope, respectively ($\xi = 0.2$). The parameters q and θ_j indicate the frequency and phase of the cosine carrier, respectively ($q = 0.25$, $\theta_1 = -0.5\pi$, and $\theta_2 = 0.5\pi$). We defined the center position o_k^1 ($k = 1, 2, \dots, 10$) as $-0.1, -2/30, -1/30, 0, 1/30, 2/30, 0.1, 0.2, 0.45$, and 0.7 so that the distribution of preferred disparities is matched to the MT data (DeAngelis & Uka, 2003). $o_k^2 = -o_k^1$ for $k = 1, 2, \dots, 10$. In the fitting procedure, original a in [Equation 4](#) was replaced with $d'(d) = \frac{a}{2} \sum_k \{G_k^2(d) - G_k^1(d)\}$ so that the decision variable is based on subtraction between the pooled response of the near population from that of far.

Acknowledgments

We thank Gregory Horwitz, Hiroki Tanaka, Hiroshi Shiozaki, and Izumi Ohzawa for helpful comments on the

manuscript, Hiroshi Shiozaki and Takanori Uka for discussion throughout the course of study, and Peter Karagianis for help in improving the English. This work was supported by grants to I.F. from the Ministry of Education, Culture, Sports, Science and Technology, Japan Science and Technology Agency (CREST) and the Takeda Science Foundation. T.D. was supported by the Japan Society for the Promotion of Science Research Fellowship for Young Researchers.

Commercial relationships: none.

Corresponding author: Ichiro Fujita.

Email: fujita@fbs.osaka-u.ac.jp.

Address: 1-3 Machikaneyama, Toyonaka, Osaka, 560-8531, Japan.

References

- Bishop, P. O., & Henry, G. H. (1971). Spatial vision. *Annual Review of Psychology*, 22, 119–160. [[PubMed](#)]
- Cogan, A. I., Lomakin, A. J., & Rossi, A. F. (1993). Depth in anticorrelated stereograms: Effects of spatial density and interocular delay. *Vision Research*, 33, 1959–1975. [[PubMed](#)]
- Cormack, L. K., Stevenson, S. B., & Schor, C. M. (1991). Interocular correlation, luminance contrast and cyclopean processing. *Vision Research*, 31, 2195–2207. [[PubMed](#)]
- Cormack, L. K., Stevenson, S. B., & Schor, C. M. (1993). Disparity-tuned channels of the human visual system. *Visual Neuroscience*, 10, 585–596. [[PubMed](#)]
- Cumming, B. G., & DeAngelis, G. C. (2001). The physiology of stereopsis. *Annual Review of Neuroscience*, 24, 203–238. [[PubMed](#)]
- Cumming, B. G., & Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature*, 389, 280–283. [[PubMed](#)]
- Cumming, B. G., & Parker, A. J. (1999). Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. *Journal of Neuroscience*, 19, 5602–5618. [[PubMed](#)] [[Article](#)]
- Cumming, B. G., Shapiro, S. E., & Parker, A. J. (1998). Disparity detection in anticorrelated stereograms. *Perception*, 27, 1367–1377. [[PubMed](#)]
- Dean, A. F. (1981). The variability of discharge of simple cells in the cat striate cortex. *Experimental Brain Research*, 44, 437–440. [[PubMed](#)]
- DeAngelis, G. C., Cumming, B. G., & Newsome, W. T. (1998). Cortical area MT and the perception of stereoscopic depth. *Nature*, 394, 677–680. [[PubMed](#)]

- DeAngelis, G. C., & Uka, T. (2003). Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *Journal of Neurophysiology*, *89*, 1094–1111. [[PubMed](#)] [[Article](#)]
- Haefner, R. M., & Cumming, B. G. (2008). Adaptation to natural binocular disparities in primate V1 explained by a generalized energy model. *Neuron*, *57*, 147–158. [[PubMed](#)]
- Janssen, P., Vogels, R., Liu, Y., & Orban, G. A. (2003). At least at the level of inferior temporal cortex, the stereo correspondence problem is solved. *Neuron*, *37*, 693–701. [[PubMed](#)]
- Jones, R. (1977). Anomalies of disparity detection in the human visual system. *Journal of Physiology*, *264*, 621–640. [[PubMed](#)] [[Article](#)]
- Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.
- Krug, K., Cumming, B. G., & Parker, A. J. (2004). Comparing perceptual signals of single V5/MT neurons in two binocular depth tasks. *Journal of Neurophysiology*, *92*, 1586–1596. [[PubMed](#)] [[Article](#)]
- Kumano, H., Tanabe, S., & Fujita, I. (2008). Spatial frequency integration for binocular correspondence in macaque area V4. *Journal of Neurophysiology*, *99*, 402–408. [[PubMed](#)] [[Article](#)]
- Lippert, J., & Wagner, H. (2001). A threshold explains modulation of neural responses to opposite-contrast stereograms. *Neuroreport*, *12*, 3205–3208. [[PubMed](#)]
- Masson, G. S., Busetini, C., & Miles, F. A. (1997). Vergence eye movements in response to binocular disparity without depth perception. *Nature*, *389*, 283–286. [[PubMed](#)]
- Neri, P. (2005). A stereoscopic look at visual cortex. *Journal of Neurophysiology*, *93*, 1823–1826. [[PubMed](#)] [[Article](#)]
- Neri, P., Parker, A. J., & Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. *Nature*, *401*, 695–698. [[PubMed](#)]
- Norcia, A. M., Sutter, E. E., & Tyler, C. W. (1985). Electrophysiological evidence for the existence of coarse and fine disparity mechanisms in human. *Vision Research*, *25*, 1603–1611. [[PubMed](#)]
- Ogle, K. N. (1952). Disparity limits of stereopsis. *Archives of Ophthalmology*, *48*, 50–60. [[PubMed](#)]
- Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1990). Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science*, *249*, 1037–1041. [[PubMed](#)]
- Orban, G. A., Janssen, P., & Vogels, R. (2006). Extracting 3D structure from disparity. *Trends in Neurosciences*, *29*, 466–473. [[PubMed](#)]
- Parker, A. J. (2007). Binocular depth perception and the cerebral cortex. *Nature Reviews Neuroscience*, *8*, 379–391. [[PubMed](#)]
- Poggio, G. F., & Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology*, *40*, 1392–1405. [[PubMed](#)]
- Poggio, G. F., Gonzalez, F., & Krause, F. (1988). Stereoscopic mechanisms in monkey visual cortex: Binocular correlation and disparity selectivity. *Journal of Neuroscience*, *8*, 4531–4550. [[PubMed](#)] [[Article](#)]
- Preston, T. J., Li, S., Kourtzi, Z., & Welchman, A. E. (2008). Multivoxel pattern selectivity for perceptually relevant binocular disparities in the human brain. *Journal of Neuroscience*, *28*, 11315–11327. [[PubMed](#)] [[Article](#)]
- Prince, S. J., & Eagle, R. A. (2000). Weighted directional energy model of human stereo correspondence. *Vision Research*, *40*, 1143–1155. [[PubMed](#)]
- Read, J. C. (2002). A Bayesian model of stereopsis depth and motion direction discrimination. *Biological Cybernetics*, *86*, 117–136. [[PubMed](#)]
- Read, J. C., & Eagle, R. A. (2000). Reversed stereo depth and motion direction with anti-correlated stimuli. *Vision Research*, *40*, 3345–3358. [[PubMed](#)]
- Read, J. C., Phillipson, G. P., Serrano-Pedraza, I., Milner, A. D., & Parker, A. J. (2010). Stereoscopic vision in the absence of the lateral occipital cortex. *PLoS One*, *5*, e12608. [[PubMed](#)] [[Article](#)]
- Rogers, B. J., & Anstis, S. M. (1975). Reversed depth from positive and negative stereograms. *Perception*, *4*, 193–201.
- Shadlen, M. N., Britten, K. H., Newsome, W. T., & Movshon, J. A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *Journal of Neuroscience*, *16*, 1486–1510. [[PubMed](#)] [[Article](#)]
- Stevenson, S. B., Cormack, L. K., & Schor, C. M. (1994). The effect of stimulus contrast and interocular correlation on disparity vergence. *Vision Research*, *34*, 383–396. [[PubMed](#)]
- Stevenson, S. B., Cormack, L. K., Schor, C. M., & Tyler, C. W. (1992). Disparity tuning in mechanisms of human stereopsis. *Vision Research*, *32*, 1685–1694. [[PubMed](#)]
- Takemura, A., Inoue, Y., Kawano, K., Quaia, C., & Miles, F. A. (2001). Single-unit activity in cortical area MST associated with disparity-vergence eye movements: Evidence for population coding. *Journal of Neurophysiology*, *85*, 2245–2266. [[PubMed](#)] [[Article](#)]
- Tanabe, S., & Cumming, B. G. (2008). Mechanisms underlying the transformation of disparity signals

- from V1 to V2 in the macaque. *Journal of Neuroscience*, 28, 11304–11314. [[PubMed](#)] [[Article](#)]
- Tanabe, S., Umeda, K., & Fujita, I. (2004). Rejection of false matches for binocular correspondence in macaque visual cortical area V4. *Journal of Neuroscience*, 24, 8170–8180. [[PubMed](#)] [[Article](#)]
- Tanabe, S., Yasuoka, S., & Fujita, I. (2008). Disparity-energy signals in perceived stereoscopic depth. *Journal of Vision*, 8(3):22, 1–10, <http://www.journalofvision.org/content/8/3/22>, doi:10.1167/8.3.22. [[PubMed](#)] [[Article](#)]
- Tyler, C. W. (1990). A stereoscopic view of visual processing streams. *Vision Research*, 30, 1877–1895. [[PubMed](#)]
- Uka, T., & DeAngelis, G. C. (2004). Contribution of area MT to stereoscopic depth perception: Choice-related response modulations reflect task strategy. *Neuron*, 42, 297–310. [[PubMed](#)]
- Uka, T., & DeAngelis, G. C. (2006). Linking neural representation to function in stereoscopic depth perception: Roles of the middle temporal area in coarse versus fine disparity discrimination. *Journal of Neuroscience*, 26, 6791–6802. [[PubMed](#)] [[Article](#)]
- Uka, T., Tanabe, S., Watanabe, M., & Fujita, I. (2005). Neural correlates of fine depth discrimination in monkey inferior temporal cortex. *Journal of Neuroscience*, 25, 10796–10802. [[PubMed](#)] [[Article](#)]
- Wilcox, L. M., & Allison, R. S. (2009). Coarse-fine dichotomies in human stereopsis. *Vision Research*, 49, 2653–2665. [[PubMed](#)]