Independent mechanisms for bright and dark image features in a stereo correspondence task

Jenny C. A. Read

Institute of Neuroscience, Newcastle University,
Newcastle upon Tyne, UK

Xavier A. Vaz

Faculty of Medical Sciences, Newcastle University,
Newcastle upon Tyne, UK

Ignacio Serrano-Pedraza

Faculty of Psychology, Complutense University of Madrid,
Madrid, Spain

A pioneering study by J. M. Harris and A. J. Parker (1995) found that disparity judgments using random-dot stereograms were better for stimuli composed of mixed bright and dark dots than when the dots were all bright or all dark. They attribute this to an improvement in stereo correspondence. This result is hard to explain within current models of how stereo correspondence is achieved. However, their experiment varied task difficulty by adding disparity noise. We wondered if this might challenge mechanisms subsequent to the solution of the correspondence problem rather than mechanisms that solve the correspondence problem itself. If so, this would avoid the need to modify current models of stereo correspondence. We therefore repeated Harris and Parker's experiment using interocular decorrelation to vary task difficulty. This technique is believed to probe stereo correspondence more specifically. We observed the efficiency increase reported by Harris and Parker for mixed-polarity dots both using their original technique of disparity noise and using interocular decorrelation. We show that this effect cannot be accounted for by the stereo energy or by simple modifications of it. Our results confirm Harris and Parker's original conclusion that mixed-polarity dots specifically benefit stereo correspondence and point up the challenge to current models of this process.

Introduction

Stereopsis is the 3D depth perception we have by virtue of seeing the world through two offset eyes. The crucial first step in this process is stereo correspondence: correctly matching up image features in the two eyes that are views of the same object. This process is believed to begin in primary visual cortex with neurons that are selectively sensitive to binocular disparity. The stereo energy model and its variants (Ohzawa, 1998; Ohzawa, DeAngelis, & Freeman, 1990, 1996, 1997; Qian & Andersen, 1997; Read, Parker, & Cumming, 2002) successfully capture many aspects of these cells' responses, including their sensitivity to disparity in random-dot stereograms, which contain no other cues to depth. Energy-model units compute a local cross-correlation between spectrally filtered image patches in each eye (Fleet, Wagner, & Heeger, 1996; Qian & Zhu, 1997). Energy- and correlation-based disparity detectors form the basis of many physiologically based models of human stereopsis (Banks, Gepshtein, & Landy, 2004; Filippini &

Banks, 2009; Fleet et al., 1996; Lippert & Wagner, 2002; Mikaelian & Qian, 2000; Prince & Eagle, 2000; Qian, 1994, 1997; Qian & Andersen, 1997; Qian & Zhu, 1997; Read, 2002a, 2002b, 2010).

However, one result in the literature challenges this commonly accepted view of disparity encoding. Harris and Parker (1995) measured statistical efficiency on a simple relative disparity task. Observers were shown a random-dot stereogram depicting two rectangular regions at different depths, separated by a vertical boundary, and asked to judge which side of the boundary was closer. Task difficulty was modulated by adding Gaussian disparity noise to each dot (Figure 1). Accurate judgments therefore required averaging over a certain number of dots on each side of the boundary, in order to average out the noise. By measuring the observer's performance, Harris and Parker could calculate the effective number of dots they were averaging over and, hence, their statistical efficiency: the number of dots actually used divided by the number of dots available in the image. Harris and Parker showed that this efficiency measure doubled when the random-dot stereogram consisted of black and white dots
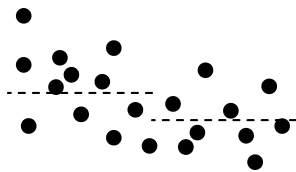
Figure 1. Disparity-noise task of Harris and Parker, top-down view. The dots have a Gaussian distribution in depth, with a mean disparity indicated by the dashed lines (disparity step stimulus). The task is to judge which side has the nearer mean disparity. Even when each dot is unambiguously located in depth, as shown here, the task is hard because of the scatter between dots.

on a gray background compared to when it consisted of only black or only white dots. They argued that this is because mixed-polarity dots reduce the complexity of the stereo correspondence problem. Random-dot stereograms offer a multitude of false matches, because each dot in the left eye could be matched with any of a large number of identical dots in the right eye. If the visual system matches black dots with black dots and white dots with white dots, the number of potential matches is halved in mixed-polarity dot patterns compared with those where all dots are identical.

This explanation implicitly invokes a "feature-matching" model of stereo correspondence, in which features such as individual dots are matched between eyes. As we show in this paper, this is hard to reconcile with a correlation-based or energy-model approach. Thus, Harris and Parker's result presents a major challenge for models of stereo correspondence based on such units.

We therefore wondered if there might be an alternative way to explain Harris and Parker's result. For example, suppose performance on their task was not limited by stereo correspondence. They made the task hard for their subjects by adding disparity noise, but this would make the task harder even if stereo correspondence remained perfect. Figure 1 shows a sketch of their stimulus, viewed top-down. Dots are scattered about the two depth planes. Each dot is depicted with a clear position in depth, and yet the task is hard because it requires the viewer to judge the mean depth of a dispersed cloud of dots. This is presumably done somewhere in the visual system after stereo correspondence has already been achieved. Suppose that this module of the visual system is only able to average over a small number of same-polarity dots but is able to average over bright and dark dots separately. This could then explain the doubling of efficiency reported by Harris and Parker, without needing to modify existing models of stereo correspondence.

This alternative explanation lacks the intuitive appeal of Harris and Parker's original technique. It seems immediately clear that black dots should only be matched with black dots and, therefore, obvious that changing to mixed black and white dots halves the complexity of stereo

correspondence. It is far less obvious why the visual system should be limited by polarity in its ability to average disparities. However, rather than debate the plausibility of this suggestion, we determined to test it experimentally.

We therefore repeated Harris and Parker's groundbreaking experiment using a stimulus designed specifically to probe the visual system's ability to acquire disparities via stereo correspondence rather than its ability to average disparities once they have been acquired. To this end, we did not add disparity noise about the depth step but increased task difficulty by reducing the binocular correlation of the stimulus. That is, we removed a certain fraction of the dots in one eye and replaced them with dots of identical luminance but at random locations in the image. The decorrelated dots were in general at different vertical locations from their original partners in the other eye and, thus, had no valid match. We also made the stimulus dynamic, displaying a completely new random-dot pattern every 150 ms (with the same disparity and binocular correlation). This effectively removes feature-matching approaches to stereo vision, since each feature in the image is immediately removed and masked by the next frame. Achieving stereo correspondence in this stimulus requires effectively identifying and discarding the unpaired dots. However, once this has been achieved, the task is trivial, since only one disparity is present on each side of the edge. Thus, if the efficiency advantage of mixed-polarity stimuli occurs after stereo correspondence, we would not expect it to improve performance on the reduced-correlation task, where performance is limited by the ability to achieve stereo correspondence. If, on the other hand, mixed-polarity dots provide an advantage at the stereo correspondence stage, as Harris and Parker proposed, we would expect the advantage to persist on the reduced-correlation task.

In this paper, we therefore compare observers' performance on the disparity-noise and reduced-correlation versions of the task and examine whether they show the same increase in efficiency when the stimulus contains both bright and dark dots compared to when it contains only bright or only dark dots. We compare the results of human observers to the properties of the stereo energy model with both linear and non-linear binocular combination.

Materials and methods

Subjects

Seven subjects aged between 19 and 38 years took part in the experiments. All subjects had experience in psychophysical experiments and only JCAR and XAV were aware of the purpose of the experiments. All subjects wore their normal visual correction (if any).
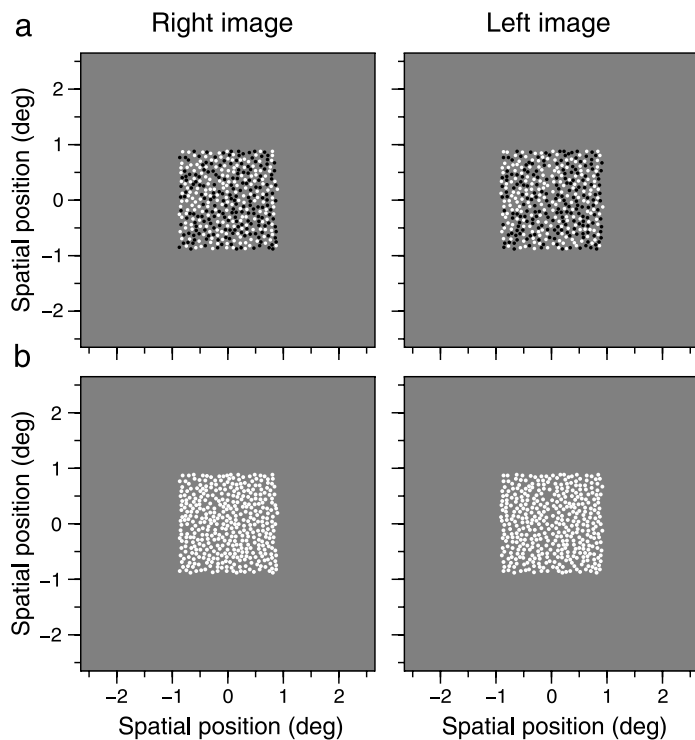
Figure 2. Sample stimuli. (a) Mixed-polarity stimulus (black and white dots). (b) Same-polarity stimulus. Stimulus parameters: N = 396 dots, no overlap, binocular correlation C = 100%, disparity step Δ = 1.4′, disparity noise σ = 3.0′ (when viewed in experimental apparatus). These stimuli can be viewed either with crossed or with divergent fusion; it will simply reverse the sign of the disparity step.

Apparatus

The experimental apparatus was as described in Serrano-Pedraza and Read (2009). Briefly, the stereo images were displayed on a rear projection screen using a passive polarization system supplied by Virtalis (Manchester, UK). The images were carefully aligned to within a pixel everywhere within the central region of the screen (where the stimuli were displayed) to ensure that as far as possible the only disparities were those introduced by the experimenter. Stimuli were generated in real time by a computer using the Windows XP operating system and a GeForce Quadro FX380 graphics card. Stimuli were programmed in Matlab 2007b (version 7.5.0.342, 32 bits; The Mathworks, www.mathworks.com) using the Psychophysics Toolbox version 3.0.8 (Brainard, 1997; Pelli, 1997) and presented on two identical DLP projectors (model FX2+ from Projection Design, Gamle Fredrikstad, Norway). The projectors were calibrated using a photometer (Minolta LS-100) to give a linear luminance response. The images were 1400 × 1050 pixels and occupied 71 × 53 cm on the projection screen. Observers viewed the screen from a distance of 160 cm in a head and

chin rest (UHCOTech HeadSpot, Houston). Each pixel therefore subtended just over 1 arcmin. Gray was 24 cd/m^2, white was 49 cd/m^2, and black was 0.7 cd/m^2.

Stimuli

Sample stimuli are shown in Figure 2 for cross-fusion. The projector screen was black apart from a central square region of 5.3 × 5.3 degrees, which was gray and contained the random-dot patterns. Random-dot patterns subtended 106 × 106 arcmin. Dots were circular with a diameter of 3.2 arcmin. Anti-aliasing was used to depict the circular dots and subpixel disparities. The number of dots was N = 396 except where stated. There were 3 contrast conditions: mixed polarity (equal numbers of black and white dots, Figure 2a), all black, and all white (Figure 2b). Observers completed runs consisting of 600 trials in total, made up of 200 all-black, 200 all-white, and 200 mixed-polarity conditions randomly interleaved. The task was always to decide which side of the stimulus, left or right, was closer to the observer. Subjects were allowed to view the stimulus for as long as they wanted; the stimulus only advanced once a response was reported via mouse button press.

The mean disparity structure of the stimulus depicted a vertical depth step, with a boundary running vertically down the center of the image and the left and right sides having equal and opposite disparities of $\pm\Delta/2$ with respect to the screen. To achieve a given binocular correlation C, a fraction $(1 - C)$ of the dots were placed at independent locations in the two eyes (still within the 106 × 106 arcmin extent of the stimulus). The remaining CN dots were placed in identical vertical locations on the screen, with horizontal disparity randomly drawn from a Gaussian distribution about the mean, with standard deviation σ. We did not slope the stimulus disparity back to zero at the edges of the screen, as Harris and Parker did, meaning that our stimulus offered a monocular cue to depth. In practice, this cue was much less helpful than the disparity cue, and the fact that we obtained the same results as Harris and Parker indicates that this difference in the stimulus was not important.

Where both black and white dots are present, overlapping dots present a problem. Allowing one dot to occlude the other presents a cue to their relative depth. If this agrees with the disparity cue, it offers a way of doing the task without using the stereo mechanisms we are trying to probe. If it conflicts with the disparity cue, it risks disturbing the stereo mechanisms we are trying to probe. Setting the luminance to the gray background where the dots overlap would be one option but risks giving the image a strange half-eaten appearance. In our "no-overlap" condition, we therefore decided to arrange the dots so that none of them overlapped, thereby sidestepping the problem. This is the solution adopted by Harris and Parker (personal communication). We also
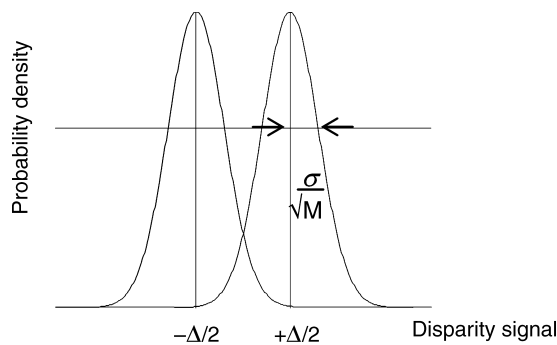
Figure 3. Probability density functions for the disparity signals on either side of the depth boundary, if the observer averages over M dots.

examined an "occlusion" condition, in which dots were simply scattered at random so that there were many places where dots overlapped. In this case, for the mixed-polarity condition, regions of overlap were given the color of whichever dot was drawn last. We used a slightly larger number of dots, 480, so that the amount of gray background was similar despite the overlap.

Stimulus generation algorithm

To generate the stimulus, then, we added dots one at a time. For each dot, we first decided at random whether it was binocularly correlated or not. If it was correlated (probability C), we picked a random cyclopean position (x, y), with x and y drawn independently with uniform probability from the range $[-W, W]$, where W was the stimulus half-width and $(0, 0)$ was the center of the stimulus. We gave the dot disparity δ, drawn at random from a Gaussian distribution with mean $\pm$ sign$(x_c) \times \Delta/2$ and standard deviation σ. The $\pm$ controls whether the left side or right side is nearer; on each trial, either $+$ or $-$ was chosen at random. We use a sign convention in which near disparity is negative. After adding this disparity, the dot's position in the left and right eyes was $(x_L, y_L) = (x_c - \delta/2, y)$ and $(x_R, y_R) = (x_c + \delta/2, y)$, respectively. If the dot was not binocularly correlated (probability $1 - C$), we drew (x_L, y_L) and (x_R, y_R) independently from $[-W, W]$.

We then looked to see whether this dot overlapped with any dots already placed in the stimulus. That is, we computed the distances between (x_L, y_L) and the centers of all dots already in place in the left eye and similarly for the right eye. If any of these distances exceeded the dot diameter, we rejected that position and generated a new pair (x_L, y_L) and (x_R, y_R). We repeated this process until a non-overlapping position had been found. For the mixed-polarity condition, we finally chose the dot's color, black or white with equal probability. This process was repeated

until all N dots had been placed in the stimulus. Figure 2 shows two sample stimuli.

Ideal observers and statistical efficiency

Following Harris and Parker (1995), we assume that an ideal observer would perform this task by accurately assigning dots to the left- or right-hand side of the depth boundary and averaging the disparity of dots on either side of the boundary. We further assume that the disparity of each dot is available with an effective noise level s, reflecting both the externally applied disparity noise σ and any internal noise. Finally, we assume that human observers manage to average only some fraction E of the available dots, where E is by definition the statistical efficiency of the observer. If the observer averages the disparity of M dots, each of which has disparity drawn from the normal distribution $N(\Delta/2, s)$, the resulting average disparity is drawn from $N(\Delta/2, s/\sqrt{M})$, where Δ is the relative disparity between the two surfaces separated by the disparity step. The distributions of this average signal on either side of the depth boundary are shown in Figure 3. We assume that the observer judges which side is closer by comparing the two numbers drawn from these distributions and assigns the closer side to be that with the smaller number. The observer's probability of getting the correct answer is then

$$P(\Delta, \sigma, M) = \frac{1}{2}\left\{1 + \text{erf}\left\{\frac{\Delta\sqrt{M}}{2s}\right\}\right\}, \qquad (1)$$

where s is an unknown function of σ and erf is the Gaussian error function or probability integral. We can rearrange this to derive the number of dots that the observer is using from each side of the depth boundary, given their performance level P:

$$M = \left[\frac{2s}{\Delta} \times \text{erf}^{-1}(2P - 1)\right]^2. \qquad (2)$$

Since on each side of the boundary there are $CN/2$ dots carrying disparity information, where C is the binocular correlation, we can then calculate the observer's efficiency as

$$E = \frac{2M}{CN} = \frac{8s^2}{CN\Delta^2}\left[\text{erf}^{-1}(2P - 1)\right]^2. \qquad (3)$$

For our purposes, the relevant quantity is the ratio of the efficiency in the mixed-contrast condition (BW) to that in
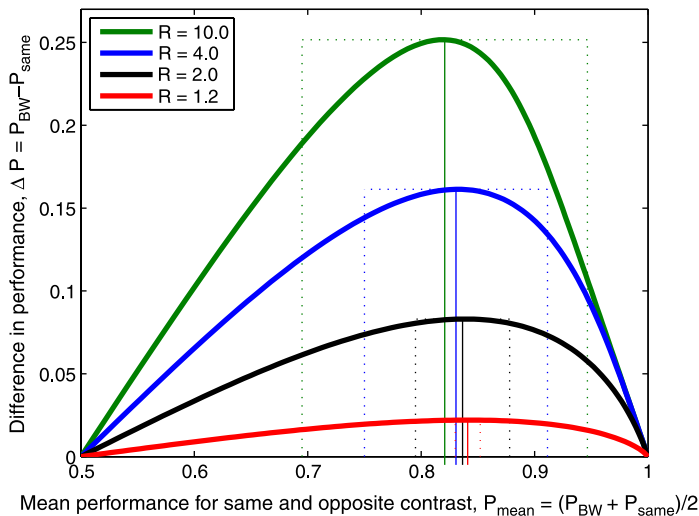
Figure 4. Predicted difference in performance for same vs. mixed polarity, for different efficiency ratios R. The solid vertical lines mark the P_{mean} at which ΔP peaks for a given R, while the dashed lines to either side mark the corresponding values of P_{same} and P_{BW}. For example, for $R = 2$, the greatest possible difference in performance is $\Delta P = 0.083$, obtained when performance is $P_{same} = 0.795$ for the same-polarity stimulus and $P_{BW} = 0.878$ for the mixed-polarity stimulus.

the same-contrast condition (B or W), comparing stimuli where the parameters dot number N, correlation C, relative disparity Δ, and disparity noise σ are all the same. If we average trials in the two same-contrast conditions, the efficiency ratio is then

$$R = \left[\frac{\text{erf}^{-1}(2P_{BW} - 1)}{\text{erf}^{-1}(P_B + P_W - 1)} \right]^2. \quad (4)$$

We used Monte Carlo resampling to estimate confidence intervals for each estimate of R. For each condition (B, W, BW), $P_{condition}$ is defined as $n_{correct}/n_{trials}$. We used the binornd function in Matlab to generate a new n_{resamp} from a binomial distribution with parameters $P_{condition}$ and n_{trials}. This produced a new estimate of $P_{condition}$, $P_{resamp} = n_{resamp}/n_{trials}$. We did this for each condition so as to arrive at a new R_{resamp}. We repeated this 10,000 times and took the 95% confidence limits to be the 2.5% and 97.5% percentiles of the resulting set of R_{resamp}.

Setting the optimal performance level

Clearly, if the task is too easy, performance will be 100% for all conditions, and we will not be able to measure the efficiency; conversely, if the task is too hard, performance will be at chance and again we will not be able to assess efficiency. If we assume that the efficiency

ratio R is constant independent of the stimulus parameters, we can ask how we should adjust task difficulty in order to maximize the difference in performance between same- and mixed-polarity conditions. Suppose that P_{same} is the proportion correct in both same-contrast conditions ($P_B = P_W = P_{same}$), and P_{BW} is the proportion correct in the mixed-contrast condition. Define $P_{mean} = (P_{BW} + P_{same}) / 2$ and $\Delta P = P_{BW} - P_{same}$. Then

$$R = \left[\frac{\text{erf}^{-1}(2P_{mean} + \Delta P - 1)}{\text{erf}^{-1}(2P_{mean} - \Delta P - 1)} \right]^2. \quad (5)$$

Figure 4 shows how the value of ΔP satisfying Equation 5 varies as a function of P_{mean}, for 4 sample values of R. Obviously, larger performance differences are possible when the efficiency advantage of mixed-polarity dots is greater. However, over a very wide range of R including the value of ~2 reported by Harris and Parker, the difference is maximized if mean performance is around 83–85%. In the experiments reported below, therefore, we tried to adjust the stimulus parameters for individual subjects so as to set mean performance at about this level. Thus, as well as the results below, each subject initially collected a small amount of pilot data at a few different difficulty levels, starting with the zero-noise, 100% correlation condition in order to familiarize them with the task.

Energy-model simulations

The simulations of disparity-selective neurons were written in Matlab (www.mathworks.com) and are available in the Supplementary materials. The images were random-dot patterns similar to those in the experiments (see Figure 11) but using square dots since this was easier to code and makes no difference to the output of the model.

We simulated the response of a population of binocular neurons tuned to a range of horizontal disparities. The left and right receptive fields were vertically oriented Gabor functions of sine ($\phi = \pi/2$) and cosine phases ($\phi = 0$), with a carrier frequency of $f = 0.025$ cycle per pixel and an envelope standard deviation of $\sigma = 10$ pixels:

$$G(f, \sigma, \phi, x_0) = \exp\left[-\frac{(x - x_0)^2 + y^2}{2\sigma^2} \right] \cos[2\pi f(x - x_0) + \phi].$$

$$(6)$$

The disparity tuning was given by a positional offset between the peak of the envelope, x_0, in the left and right eyes.
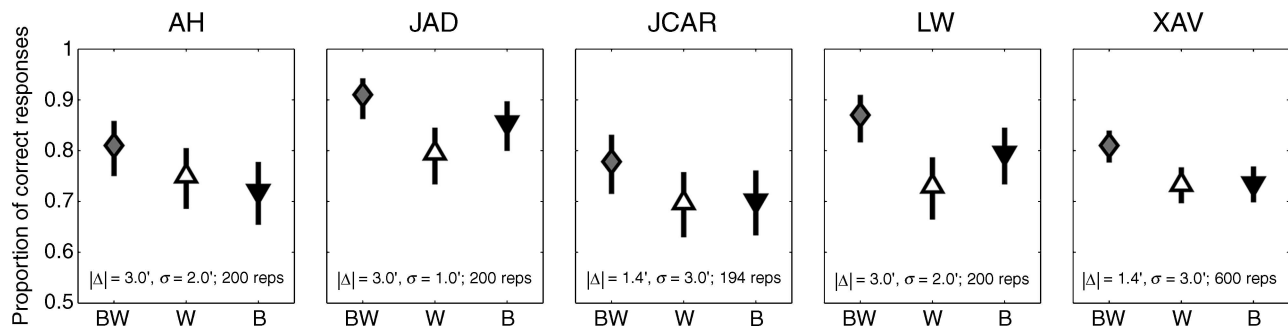
Figure 5. Performance on the disparity-noise task with no dot overlap, for 5 different subjects. Symbols show performance in the 3 different conditions; error bars show 95% confidence intervals assuming simple binomial statistics. The code above each panel identifies the subject. Stimulus parameters: $N = 396$ dots, no overlap, binocular correlation $C = 100\%$. Disparity step Δ and disparity noise σ were adjusted for each subject individually, and subjects performed at least 200 repetitions of each condition, as indicated in each panel.

We computed the inner product of the left- and right-eye images, $I_L(x, y)$ and $I_R(x, y)$, with odd and even receptive fields:

$$v_{LE}(\Delta x) = \iint dx\,dy\,G\left(f, \sigma, 0, -\frac{\Delta x}{2}\right) I_L(x, y),$$

$$v_{LO}(\Delta x) = \iint dx\,dy\,G\left(f, \sigma, \frac{\pi}{2}, -\frac{\Delta x}{2}\right) I_L(x, y),$$

$$v_{RE}(\Delta x) = \iint dx\,dy\,G\left(f, \sigma, 0, \frac{\Delta x}{2}\right) I_R(x, y),$$

$$v_{RO}(\Delta x) = \iint dx\,dy\,G\left(f, \sigma, \frac{\pi}{2}, \frac{\Delta x}{2}\right) I_R(x, y). \tag{7}$$

In this paper, we present results for 4 different types of disparity-tuned neurons. We include both "tuned-excitatory" and "near"-type neurons (Poggio & Fischer, 1977; Read & Cumming, 2004), i.e., cells whose disparity tuning function is symmetric about a central peak and those with one peak and one trough. We include both complex cells, whose response is independent of stimulus phase, and simple cells, whose response depends on phase. Finally, we use both the original energy model of Ohzawa et al. (1990), in which monocular inputs are combined linearly ("ODF"), and the modified version proposed by Read et al. (2002), in which monocular inputs are thresholded before being combined ("RPC"). The response of these four types is given as follows:

i) A tuned-excitatory ODF complex cell:
   $C_{ODF:TE}(\Delta X) = (v_{LE} + v_{RE})^2 + (v_{LO} + v_{RO})^2$
ii) A near ODF simple cell:
   $C_{ODF:NE}(\Delta X) = (v_{LO} - v_{RE})^2$
iii) A tuned-excitatory RPC complex cell:
   $C_{RPC:TE}(\Delta X) = ([v_{LE}] + [v_{RE}])^2 + ([v_{LO}] + [v_{RO}])^2$
iv) A near RPC simple cell:
   $C_{RPC:TE}(\Delta X) = ([v_{LO}] - [v_{RE}])^2.$

Results

Performance of human observers
The efficiency advantage of mixed-polarity stimulus is destroyed by occlusion

We first replicated Harris and Parker's (1995) original results. We kept the disparity noise σ constant at 3 arcmin as used by Harris and Parker and varied the size of the disparity step so as to obtain average performance at around 0.8 correct.

Figure 5 shows results for 5 subjects on the stimulus designed to be as close as possible to that of Harris and Parker. The symbols show the proportion correct, and the error bars show the 95% score confidence intervals assuming simple binomial statistics (Agresti & Coull, 1998). As reported by Harris and Parker, performance is not significantly different for all-white versus all-black dots but is significantly better for mixed-polarity stimuli, i.e., those containing equal numbers of black and white dots. Figure 6 presents these results as efficiency ratios, using Equation 4. For each subject, the efficiency ratio is significantly above 1 and is around 2, consistent with Harris and Parker's report. This confirms that we are able to replicate their results with our subjects and apparatus.

The lack of dot overlap is crucial. Figure 7 shows performance of three subjects on the same disparity-noise task but this time with dots scattered at random and allowed to overlap. This time, there was no significant improvement for the mixed-polarity condition. For 2/3 subjects, the efficiency ratios are not significantly different from 1, while the remaining subject MBAC actually performed worse with mixed black and white dots than when the dots were all black or all white. We suggest that this may be because the regions of overlap provide an occlusion cue to relative distance that sometimes agrees and sometimes conflicts with the cue provided by binocular disparity. Apparently, this conflict is enough to destroy the advantage offered by mixed-polarity stimuli in
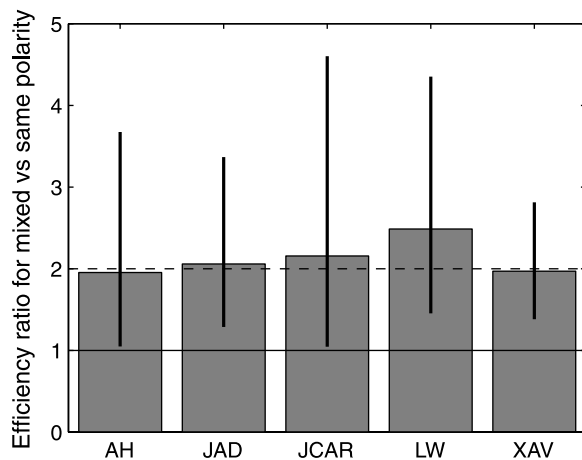
Figure 6. Efficiency ratios for the disparity-noise task with no dot overlap. Error bars are 95% confidence intervals, estimated by resampling as described in the Materials and methods section. The solid line marks an efficiency ratio of 1 (no advantage for either condition), and the dashed line marks an efficiency ratio of 2 (efficiency twice as good for mixed vs. same polarity). Stimulus parameters: N = 396 dots, no overlap, binocular correlation C = 100%. Disparity step Δ and disparity noise σ were adjusted for each subject individually, and subjects performed at least 200 repetitions of each condition, as indicated in Figure 5.

exploiting binocular disparity and, for some subjects, make them less useful than same-polarity stimuli. Harris and Parker also avoided dot overlap in their stimuli.

Similar results are obtained when decorrelation is used instead of disparity noise

Having replicated Harris and Parker's results in our apparatus, we then performed a different version of their experiment, in which task difficulty was increased not by adding disparity noise (so now σ = 0) but by reducing the interocular correlation (so now C < 1). We also made the stimulus dynamic, changing every 150 ms. As outlined above, we argued that this stimulus might challenge stereo

correspondence more specifically than the disparity-noise stimulus, which is still difficult even when correspondence is perfect. We therefore felt it important to verify that mixed-polarity stimuli continue to have an advantage in this configuration.

Figure 8 shows the results of this experiment for 4 subjects, and Figure 9 shows the resulting efficiency ratios. In 3/4 subjects, the efficiency ratio is still significantly above 1; indeed, the mean over all 4 subjects is now 3.7. For subject JAD, the estimated efficiency ratio was 1, but this may be because we did not manage to select the best correlation for this subject: his performance was close to chance, and hence, the error on our estimate of his efficiency ratio is large (95% confidence interval ranges from 0.2 to 3.6). Overall, however, it is clear that the advantage of mixed-polarity dots persists with the new stimulus. Indeed, for most subjects, moving from a static stimulus with disparity noise to a dynamic stimulus with decorrelation has enhanced the performance advantage of mixed-polarity dots.

For subject XAV, we also redid both experiments with fewer dots (N = 100). The efficiency ratios are shown in Figure 10. The efficiency ratios are somewhat reduced for the smaller dot number but are still well above 1, especially in the decorrelated stimulus.

Response of energy-model units for same- vs. mixed-polarity stimuli

Overall, then, it is clear that we have replicated Harris and Parker's finding of improved performance with mixed black and white dots, not only with their original disparity-noise stimulus but also with our decorrelated stimulus. In the Introduction section, we stated that this improved performance is puzzling given the properties of the stereo energy model. We now examine the performance of various models to these stimuli, as shown in Figure 11. We used 4 different types of model. First, we considered the original "ODF" energy model introduced

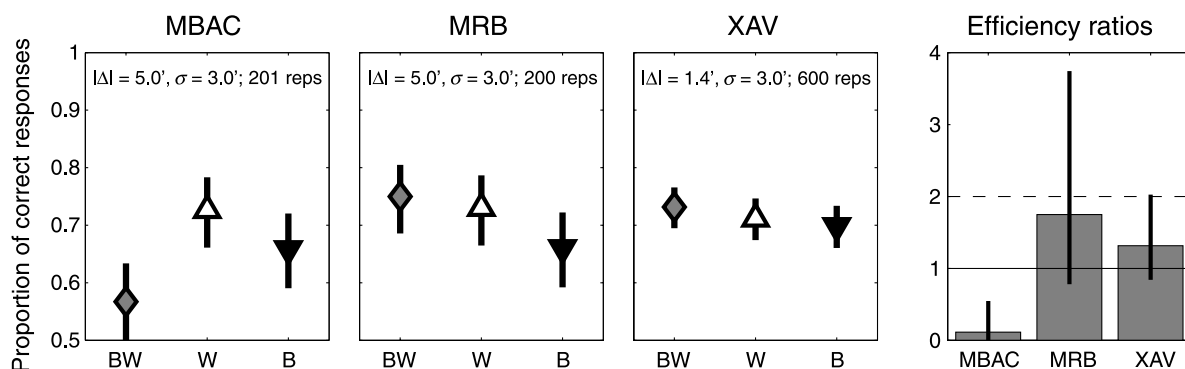Figure 7. Performance on the disparity-noise task with occlusion, for 3 different subjects. As Figure 5 except for stimulus parameters: N = 480 dots, randomly scattered with overlap. Last panel shows efficiency ratios, as in Figure 6.
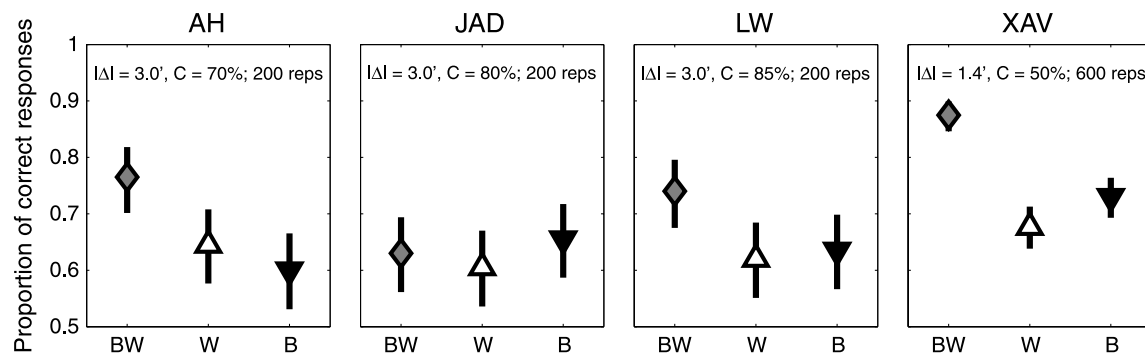
Figure 8. Performance on the decorrelation task with no dot overlap, for 4 different subjects. As Figure 5 except for stimulus parameters: N = 480 dots, not overlapping, disparity noise σ = 0. Disparity step Δ and binocular correlation C were adjusted for each subject individually, and subjects performed at least 200 repetitions of each condition, as indicated in each panel.

by Ohzawa et al. (1990), with pure position disparity ("ODF TE") and with phase disparity ("ODF Near"). This effectively implements cross-correlation between a filtered version of the left- and right-eye images. Second, we considered the "RPC" variant introduced by Read et al. (2002). This contains a non-linearity prior to binocular combination, which means that it no longer implements simple cross-correlation. The threshold non-linearity means that this model responds differently to bright and dark features, unlike the energy model that was designed as a model of complex cells and that responds equally to bright and dark features of equal contrast.

Figure 11 shows the mean population response of these 4 different types of neuronal models to 10,000 random-dot patterns with either mixed black-and-white dots (top row)

or all-white dots (bottom row). The results with all-black dots are the same as with all-white dots, so they are not shown here. The stimuli all had a uniform disparity of 10 pixels. In the top row, the mean population response modulates strongly as a function of preferred disparity. For the tuned-excitatory neurons, which are the simplest to decode, the stimulus disparity can be immediately read off from the location of the peak in the population. In the bottom row, however, this modulation is much weaker. The shape is still the same, but the amplitude of modulation is a much smaller fraction of the baseline. If we assume constant noise in both cases, this would make any readout more error-prone for same-polarity dots. It is tempting to conclude that this property of energy-like models is the neuronal basis for the impaired performance of human observers on same-polarity versus mixed-polarity dot patterns.

However, this would be premature. The reduced amplitude in Figure 11 is a side effect of the way luminance was represented in the two eyes' images, I_{L}
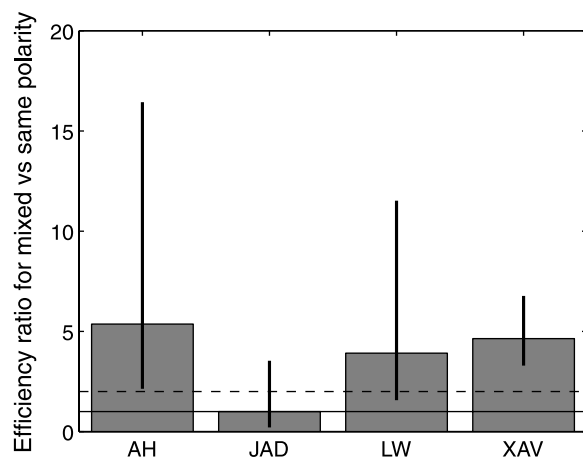

Figure 9. Efficiency ratios for the decorrelation task with no dot overlap. The solid line marks an efficiency ratio of 1 (no advantage for either condition), and the dashed line marks an efficiency ratio of 2 (efficiency twice as good for mixed vs. same polarity). Error bars are 95% confidence intervals. Stimulus parameters: N = 480 dots, not overlapping, disparity noise σ = 0. Disparity step Δ and binocular correlation C were adjusted for each subject individually, and subjects performed at least 200 repetitions of each condition, as indicated in Figure 8.
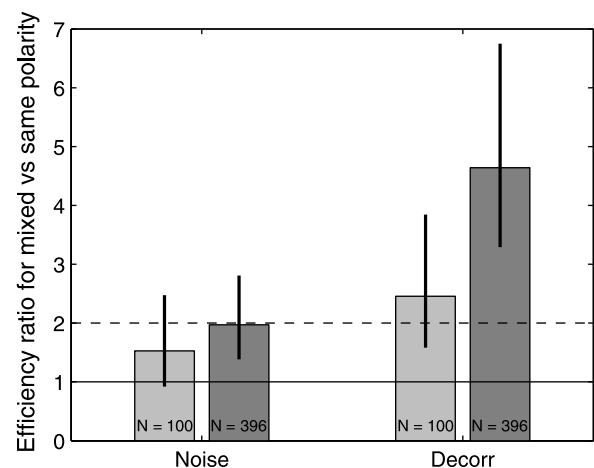

Figure 10. Efficiency ratios for subject XAV, for the noise and decorrelation tasks with N = 100 dots (light bars) and N = 396 (dark bars). Stimulus parameters: $|\Delta|$ = 1.4′ throughout; for "Noise", σ = 3′ and C = 100%; for "Decorr", σ = 0′ and C = 50%.
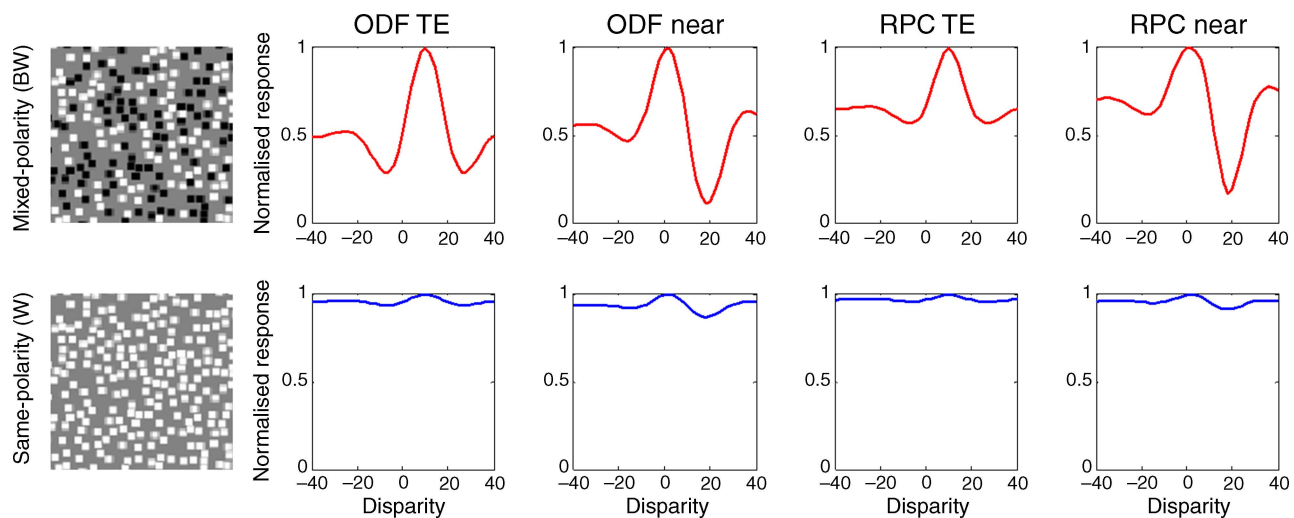
Figure 11. Mean population response of model neurons to random-dot patterns with mixed-polarity (top, red curves) and same-polarity (bottom, blue curves) dots. The stimuli had a uniform disparity of 10 pixels, while the preferred disparity of the neurons (i.e., the positional offset between the left and right receptive field envelopes) is shown on the horizontal axis. The curves show the mean response to 10,000 randomly generated dot patterns, normalized to 1. In this version, the same-polarity (white) images had much higher mean value than the mixed-polarity images. We have not shown tuning curves for images with all-black dots, since these are exactly the same as the all-white dot results.

and I_R (Equation 7). In Figure 11, the gray background was represented by 0, white dots by +1, and black dots by −1. This means that the mixed-polarity images have near-zero DC component and thus that their autocorrelation is zero for large offsets, as is usually assumed for random-dot stereograms (Prince, Pointon, Cumming, & Parker, 2002; Read & Cumming, 2003; Read et al., 2002). This in turn means that the amplitude of the modulation in firing rate as a function of disparity is equal to the baseline response, defined as the mean response to disparities far from the preferred disparity or equivalently to binocularly uncorrelated stimuli (strictly, it is equal to the baseline only for tuned-excitatory or very narrow-band cells; the

ratio is slightly less than 1 for odd-symmetric finite-bandwidth cells). For the same-contrast stimuli in Figure 11, the image functions were positive or zero everywhere. Their DC component and, hence, autocorrelation were always positive. This means that the baseline response greatly increases relative to the amplitude of modulation. As we show in Appendix A, the ratio of amplitude to baseline is always maximized if the images have zero DC component. Non-zero DC component, either positive or negative, reduces the amplitude. The amplitude measured with all-white or all-black dots is therefore lower than when measured with mixed black-and-white dots.
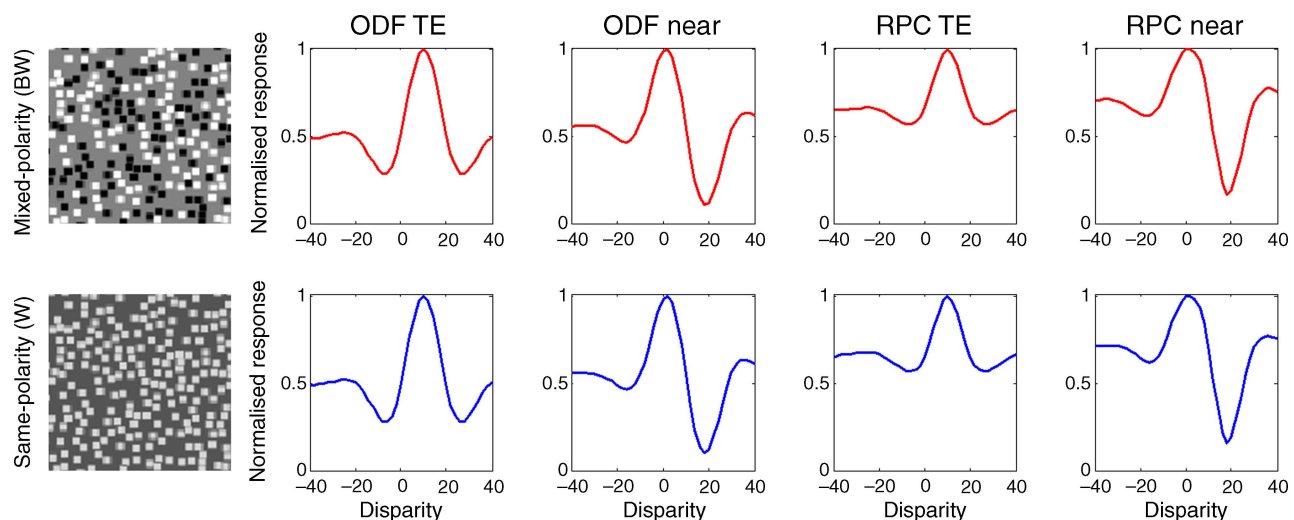

Figure 12. As Figure 11, except that both mixed- and same-polarity images had the same mean value.
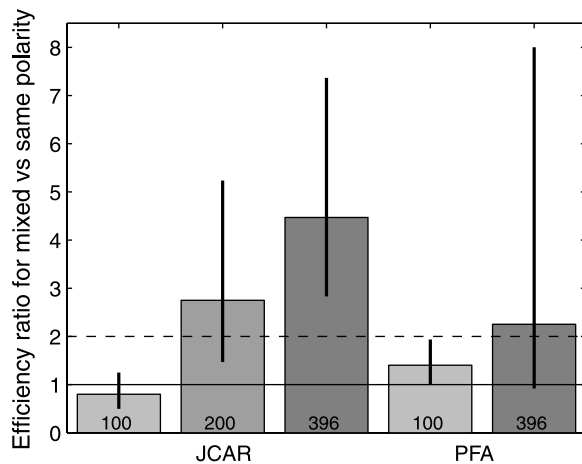
Figure 13. Efficiency ratios for two subjects, for the disparity-noise task with constant mean luminance. Stimulus parameters: N as indicated in each bar; C = 100%, σ = 2′; for JCAR, $|\Delta|$ = 2′ throughout; for PFA, $|\Delta|$ = 2′ for N = 100 and $|\Delta|$ = 1.4′ for N = 396.

Figure 12 shows the results of simulations that are identical except that the DC components of the images have been removed. That is, the gray background in the all-white dot patterns has been made slightly darker than in the black-and-white patterns, such that the mean luminance is zero in both cases. Now, the model neurons show no difference in the amplitude of their modulation, regardless of whether the images contain mixed- or same-polarity dots. This is so both for the original energy model and for a modified version of it in which the response from each eye is thresholded prior to binocular combination (Read et al., 2002). Again, the results for all-black dots are the same as for all-white dots and so are not shown here.

The responses shown in Figure 12 suggest that performance should be equal for mixed-polarity and same-polarity dots. Admittedly, by normalizing all the curves to the same peak value, we have concealed the fact that the raw numbers are larger for the mixed-polarity case, because of its larger contrast. However, this effect is unlikely to explain the better human performance for mixed-polarity stimuli. It would predict that performance would be better for brighter white dots than for dimmer white dots. In fact, stereoacuity is generally not very sensitive to small variations in contrast well above threshold, and Harris and Parker controlled for this in their original experiments.

Of course, in reality, the luminance of the images is non-zero everywhere. There is nothing intrinsically special about the gray level of the background, except insofar as it represents the average luminance of the image. Most workers in the field either tacitly or explicitly assume that the encoding of disparity effectively removes this DC component, either across the whole image or on a more local scale, when computing the match between left- and right-eye images. If this assumption is incorrect, the amplitude of disparity modulation could be reduced for same-polarity relative to mixed-polarity stimuli. For example, suppose that preprocessing prior to the disparity computation effectively subtracts off the gray background of the screen, not the mean luminance of the particular images. This would give the results shown in Figure 11.

It seems very unlikely that this effect underlies Harris and Parker's results. Certainly if it did, it would clearly be a completely different phenomenon from the one Harris and Parker claimed to have found. It does not depend on matching like-polarity features and is oblivious to whether the stimulus contains same-polarity or mixed-polarity dots, simply reflecting the mean luminance of the stimulus. Harris and Parker carefully examined whether their effect could be an artifact of changes in luminance or contrast. They reported that efficiency for their observers did not vary over a wide range of dot luminance levels. However, perhaps the simplest way to test this explanation is to measure human performance with the stimuli of Figure 12, i.e., keeping the mean luminance constant, and see if this removes the advantage of the mixed-polarity stimulus.

Results from two observers are shown in Figure 13. The stimulus was the disparity-noise stimulus of Harris and Parker (1995), with 100% correlation. In these results, the three polarity conditions were tested in separate blocks rather than interleaved. The gray background was varied so that the mean luminance across the whole stimulus was constant, as described above for Figure 12. For the energy model, this manipulation has a major effect, as shown in Figures 11 and 12, abolishing the advantage of mixed-polarity dots. For our human observers, it has no discernible effect. The advantage of mixed-polarity dots is still strong for large dot numbers and weaker for lower dot numbers, as above.

Discussion

In 1995, Harris and Parker demonstrated an advantage for mixed-polarity dots over same-polarity dots in a depth task. The efficiency gain was around 2, suggesting that twice as many dots are used in the mixed-polarity condition as in the same condition. Harris and Parker attributed this to improved stereo correspondence in the mixed-polarity condition, concluding that stereo correspondence proceeds in separate "bright" and "dark" channels.

This hypothesis is hard to reconcile with the successful and widely used stereo energy model of disparity-selective neurons or with more abstract variants such as correlation-based matching. Such models do not pay attention to image features, and it is hard to see how to build separate "bright" and "dark" channels into them.

Read, Vaz, & Serrano-Pedraza

Previous models have postulated an output non-linearity before binocular combination (Read & Cumming, 2003; Read et al., 2002; Tanaka & Ohzawa, 2006), but our simulations show that this modification has no effect on the response to same- versus mixed-polarity patterns. It is certainly easy to build implementations of the energy model that respond more strongly to mixed-polarity than to same-polarity dot patterns, but on closer examination, this property turns out to be due to confounding differences in the image contrast or mean luminance, not to the polarity of the image features per se. Control experiments demonstrate that Harris and Parker's effect does not stem from differences in overall image contrast or luminance.

Given this, we wondered if we could place Harris and Parker's separate neuronal channels at a later stage of visual processing. We argued that their disparity-noise task may not challenge stereo correspondence so much as subsequent depth processing. Even if the disparity of each dot is correctly identified, the scatter in disparities means that it is still challenging to compute the sign of the depth boundary. This task is presumably computed by higher visual areas responsible for depth perception. Perhaps these higher areas contain separate neuronal channels for bright and dark image features. If so, this would mean we could keep the existing energy model as a description of disparity encoding in early visual cortex, avoiding the need to modify it so as to include bright and dark channels.

To test this suggestion, we redid Harris and Parker's task using a slightly different stimulus, designed to specifically challenge stereo correspondence. We made the task hard by reducing the correlation between the two eyes. Many dots had no valid match in the other eye, but all matching dot pairs depicted the same magnitude of disparity. Once valid matches have been identified and the other dots discarded, the task is trivial. We hypothesized that with this stimulus, the advantage of mixed-polarity dots might vanish.

However, this hypothesis was comprehensively disproven. For most subjects, the advantage of mixed-polarity dots not only persisted with the decorrelation stimulus but also actually increased: the mean efficiency ratio, averaged across subjects, rose from 2 to nearly 4. If anything, this suggests that the advantage of mixed-polarity dots may arise at the initial stereo correspondence phase more than from the subsequent depth processing.

Our results do not completely support Harris and Parker's conclusions. They found an efficiency ratio of close to 2 throughout, largely independent of dot number, and this led them to postulate 2 independent channels. The larger efficiency ratios we find with our new stimulus (Figure 9) do not fit so well with this proposal. In addition, while we have not systematically investigated the effect of dot number, we do seem to find a stronger dependence on dot number than Harris and Parker, with efficiency ratios falling below 2 when the stimulus contains 100 dots (Figures 10 and 13).

Harris and Parker's (1995) work is just one of several studies examining whether ON and OFF pathways are pooled or processed separately in different aspects of visual processing. When attention can be directed separately to bright and dark elements, the two pathways are generally processed separately. So, for example, Hibbard, Bradshaw, and Eagle (2000) compared Dmax, the maximum step size to detect apparent motion, and found that it was limited by the number of bright or dark dots rather than the total number of stimuli. Similarly on a global motion coherence task, Croner and Albright (1997) found that motion coherence thresholds are improved if the signal dots have a different contrast polarity from the noise dots. On tasks that are not helped by feature-based attention directed to one or the other polarity, the picture that has emerged is that ON and OFF pathways are processed separately when the task involves extracting form or segregating objects in some way (Badcock, Clifford, & Khuu, 2005; Brooks & van der Zwan, 2002; Wenderoth, 1996; Wilson, Switkes, & De Valois, 2004) and pooled when the task involves extracting global properties (Edwards, 2009; Edwards & Badcock, 1994; Snowden & Edmunds, 1999).

So, for example, Edwards and Badcock (1994) found no difference in global motion coherence thresholds for same-polarity stimuli consisting of 100 bright dots versus mixed-polarity stimuli consisting of 50 bright and 50 dark dots, in which the signal dots were all light. In these stimuli, at threshold, only a small number of light dots carry the signal and most light dots are moving incoherently (as are all the bright dots). The signal does not, therefore, pop out even if attention is directed solely to the bright dots. Under these circumstances, mixed-polarity stimuli had no advantage. That is, motion with 14 bright signal dots, 36 bright noise dots, and 50 dark noise dots was no more visible than motion with 14 bright signal dots and 86 bright noise dots, indicating that the relative signal-to-noise ratio is that of all dots together (i.e., 16% in both cases) rather than in ON pathway alone (39% in the former case vs. 16% in the second). Thus, in this global motion detection task, information from the ON and OFF pathways appears to be pooled, as it is in the motion energy models on which the stereo energy model was based (Adelson & Bergen, 1985, 1986; Watson & Ahumada, 1985). Conversely, in a form-based motion task where observers had to track a square defined by four moving dots, Edwards (2009) found evidence that ON and OFF pathways are processed separately. Observers could tolerate higher numbers of noise dots if the noise dots were mixed polarity compared to when they were all the same polarity as the signal dots defining the moving square. Edwards argued that ON and OFF pathways are processed separately for processing form and pooled for processing global information. He argued that Harris and Parker's result reflects the use of stereo information in object grouping and segmentation. If so, this raises the question of whether a different stereo task—say, a global

motion-in-depth task—might produce a different result. However, the distinction between global motion and form processing is not clear-cut; for example, Bell and Badcock (2008) found that bright and dark information appears to be pooled in the detection of radial frequency contours, even though this is a form-detection task, and thus, ON and OFF channels might have been expected to make separate contributions.

<h1 style="background:#4a90d9;color:white;display:inline-block;padding:4px 12px;">Appendix A</h1>

The amplitude of disparity tuning is maximized when the images' DC component is removed by preprocessing

In this appendix, we justify our claim that the amplitude of disparity tuning in stereo energy-model units is always maximized when the images supplied to the model have zero DC component. Our starting point is the response of an energy-model simple cell:

$$C = v_L^2 + v_L^2 + 2v_L v_R, \tag{A1}$$

where v_L and v_R are the inner product of each image with that eye's receptive field, as in Equation 7. To compute the average response of this unit over many random images, first consider the average of the last term:

$$\langle v_L v_R \rangle = \iint dx dy \iint dx' dy' \rho_L(x,y) \rho_R(x',y')$$
$$\cdot \langle I_L(x,y) I_R(x',y') \rangle, \tag{A2}$$

where ρ_L and ρ_R represent the receptive fields (for this proof, they need not be Gabor functions). For uniform disparity stimuli in which the left and right images are identical apart from a horizontal offset of Δx, this becomes

$$\langle v_L v_R \rangle (\Delta x) = \iint dx dy \iint dx' dy' \rho_L(x,y) \rho_R(x',y')$$
$$\cdot \langle I(x,y) I(x' - \Delta x, y') \rangle. \tag{A3}$$

Suppose that the mean luminance of the random images is μ. That is, we can write

$$I(x,y) = \mu + \epsilon(x,y), \tag{A4}$$

where ε is a random variable, picked independently for each x and y from a distribution with zero mean. For

example, for a random-dot pattern with black and white dots, μ represents the luminance of the background and ε has three peaks: a peak at 0 for background pixels and symmetric peaks on either side of zero for the black and white dots. For an all-white dot pattern with more background pixels than dots, μ is slightly higher than the luminance of the background, and ε has two peaks arranged asymmetrically about 0: a small peak at a positive value, representing the white dots, and a larger peak at a negative value closer to zero, representing the gray background.

For non-corresponding points in the images (i.e., $x \neq x' - \Delta x$ or $y \neq y'$), the values of ε are uncorrelated, and so the product of the images averages to μ^2. For corresponding points (i.e., where $x = x' - \Delta x$ and $y = y'$), the values of ε are identical, and so there we pick up an additional term that depends on the variance of ε:

$$\langle I(x,y) I(x' - \Delta x, y') \rangle = \langle (+\epsilon)(+\epsilon') \rangle = \mu^2$$
$$+ \langle \epsilon^2 \rangle \delta_{\text{Dirac}}(x - x' + \Delta x, y - y'). \tag{A5}$$

Thus,

$$\langle v_L v_R \rangle (\Delta x) = \mu^2 \iint dx dy \rho_L(x,y) \iint dx' dy' \rho_R(x',y')$$
$$+ \langle \epsilon^2 \rangle \iint dx dy \rho_L(x,y) \rho_R(x + \Delta x, y), \tag{A6}$$

and similarly

$$\langle v_L^2 \rangle = \mu^2 \left[\iint dx dy \rho_L(x,y) \right]^2 + \langle \epsilon^2 \rangle \iint dx dy \rho_L^2(x,y), \tag{A7}$$

for $\langle v_R^2 \rangle$. Using these results, we can write the mean energy-model response as

$$\langle E \rangle = \mu^2 (L+R)^2 + \langle \epsilon^2 \rangle [M + B(\Delta x)], \tag{A8}$$

where

$$L = \iint dx dy \rho_L(x,y), \quad R = \iint dx dy \rho_R(x,y),$$
$$M = \iint dx dy \left[\rho_L^2(x,y) + \rho_L^2(x,y) \right],$$
$$B(\Delta x) = 2 \iint dx dy \rho_L(x,y) \rho_R(x + \Delta x, y).$$

$$\tag{A9}$$

From Equation A8, we can divide this mean response into a baseline response that would be observed even with binocularly uncorrelated stimuli

$$U = \mu^2(L+R)^2 + \langle \epsilon^2 \rangle M \qquad (A10)$$

and a disparity-modulated term $\langle \varepsilon^2 \rangle B(\Delta x)$. In the uninteresting case where the images are blank, $\langle \varepsilon^2 \rangle = 0$ and so there is no disparity modulation. Otherwise, the amplitude of the disparity tuning curve relative to the baseline is

$$A = \frac{B(\Delta x_{\text{pref}})}{M + \frac{(L+R)^2 \mu^2}{\langle \epsilon^2 \rangle}}, \qquad (A11)$$

where Δx_{pref} is defined as the disparity that maximizes the magnitude of the disparity-modulated term. *L, R, M,* and Δx_{pref} all depend only on the particular receptive field functions, i.e., the properties of the neuronal population encoding disparity. The only term that depends on the image statistics is $\mu^2/\langle \varepsilon^2 \rangle$. This term is multiplied by $(L + R)$, the integral of the receptive field functions. For the special case of odd-symmetric or very narrow-band cells, this integral is zero. In this case, the amplitude ceases to depend on the image statistics and is simply $A = B(\Delta_{\text{pref}})/M$. Where the integral $(L + R)$ is non-zero, it is clear by inspecting Equation A11 that A is maximized when the image has no DC component, i.e., $\mu = 0$. Then, $A = B(\Delta_{\text{pref}})/M$. Any non-zero value of μ reduces A, the amplitude of the disparity-modulated response. This is the reason for the difference between the mixed- and same-polarity stimuli in Figure 11.

Acknowledgments

Commercial relationships: none.
Corresponding author: Jenny C. A. Read.
Email: J.C.A.Read@ncl.ac.uk.
Address: Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, 2,* 284–299.

Adelson, E. H., & Bergen, J. R. (1986, May 7–9). *The extraction of spatio-temporal energy in human and machine vision.* Paper presented at the Workshop on Motion: Representation and Analysis, Charleston, SC.

Agresti, A., & Coull, B. (1998). Approximate is better than "exact" for interval estimation of binomial proportions. *American Statistician, 52,* 119–126.

Badcock, D. R., Clifford, C. W., & Khuu, S. K. (2005). Interactions between luminance and contrast signals in global form detection. *Vision Research, 45,* 881–889.

Banks, M. S., Gepshtein, S., & Landy, M. S. (2004). Why is spatial stereoresolution so low? *Journal of Neuroscience, 24,* 2077–2089.

Bell, J., & Badcock, D. R. (2008). Luminance and contrast cues are integrated in global shape detection with contours. *Vision Research, 48,* 2336–2344.

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

Brooks, A., & van der Zwan, R. (2002). The role of ON- and OFF-channel processing in the detection of bilateral symmetry. *Perception, 31,* 1061–1072.

Croner, L. J., & Albright, T. D. (1997). Image segmentation enhances discrimination of motion in visual noise. *Vision Research, 37,* 1415–1427.

Edwards, M. (2009). Common-fate motion processing: Interaction of the On and Off pathways. *Vision Research, 49,* 429–438.

Edwards, M., & Badcock, D. R. (1994). Global motion perception: Interaction of the ON and OFF pathways. *Vision Research, 34,* 2849–2858.

Filippini, H. R., & Banks, M. S. (2009). Limits of stereopsis explained by local cross-correlation. *Journal of Vision, 9*(1):8, 1–18, http://www.journalofvision.org/content/9/1/8, doi:10.1167/9.1.8. [PubMed] [Article]

Fleet, D., Wagner, H., & Heeger, D. (1996). Neural encoding of binocular disparity: Energy models, position shifts and phase shifts. *Vision Research, 36,* 1839–1857.

Harris, J. M., & Parker, A. J. (1995). Independent neural mechanisms for bright and dark information in binocular stereopsis. *Nature, 374,* 808–811.

Hibbard, P. B., Bradshaw, M. F., & Eagle, R. A. (2000). Cue combination in the motion correspondence problem. *Proceedings of the Royal Society B: Biological Sciences, 267,* 1369–1374.

Lippert, J., & Wagner, H. (2002). Visual depth encoding in populations of neurons with localized receptive fields. *Biological Cybernetics, 87,* 249–261.

Mikaelian, S., & Qian, N. (2000). A physiologically-based explanation of disparity attraction and repulsion. *Vision Research, 40,* 2999–3016.

Ohzawa, I. (1998). Mechanisms of stereoscopic vision: The disparity energy model. *Current Opinion in Neurobiology, 8,* 509–515.

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.

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1996). Encoding of binocular disparity by simple cells in the cat's visual cortex. *Journal of Neurophysiology, 75,* 1779–1805.

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1997). Encoding of binocular disparity by complex cells in the cat's visual cortex. *Journal of Neurophysiology, 77,* 2879–2909.

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

Poggio, G. F., & Fischer, B. (1977). Binocular interaction and depth sensitivity of striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology, 40,* 1392–1405.

Prince, S. J., Pointon, A. D., Cumming, B. G., & Parker, A. J. (2002). Quantitative analysis of the responses of V1 neurons to horizontal disparity in dynamic random-dot stereograms. *Journal of Neurophysiology, 87,* 191–208.

Prince, S. J. P., & Eagle, R. E. (2000). Weighted directional energy model of human stereo correspondence. *Vision Research, 40,* 1143–1155.

Qian, N. (1994). Computing stereo disparity and motion with known binocular cell properties. *Neural Computation, 6,* 390–404.

Qian, N. (1997). Binocular disparity and the perception of depth. *Neuron, 18,* 359–368.

Qian, N., & Andersen, R. A. (1997). A physiological model for motion-stereo integration and a unified explanation of Pulfrich-like phenomena. *Vision Research, 37,* 1683–1698.

Qian, N., & Zhu, Y. (1997). Physiological computation of binocular disparity. *Vision Research, 37,* 1811–1827.

Read, J. C. A. (2002a). A Bayesian approach to the stereo correspondence problem. *Neural Computation, 14,* 1371–1392.

Read, J. C. A. (2002b). A Bayesian model of stereopsis depth and motion direction discrimination. *Biological Cybernetics, 86,* 117–136.

Read, J. C. A. (2010). Vertical binocular disparity is encoded implicitly within a model neuronal population tuned to horizontal disparity and orientation. *PLoS Computational Biology, 6,* e1000754.

Read, J. C. A., & Cumming, B. G. (2003). Testing quantitative models of binocular disparity selectivity in primary visual cortex. *Journal of Neurophysiology, 90,* 2795–2817.

Read, J. C. A., & Cumming, B. G. (2004). Ocular dominance predicts neither strength nor class of disparity selectivity with random-dot stimuli in primate V1. *Journal of Neurophysiology, 91,* 1271–1281.

Read, J. C. A., Parker, A. J., & Cumming, B. G. (2002). A simple model accounts for the reduced response of disparity-tuned V1 neurons to anti-correlated images. *Visual Neuroscience, 19,* 735–753.

Serrano-Pedraza, I., & Read, J. C. A. (2009). Stereo vision requires an explicit encoding of vertical disparity. *Journal of Vision, 9*(4):3, 1–13, http://www.journalofvision.org/content/9/4/3, doi:10.1167/9.4.3. [PubMed] [Article]

Snowden, R. J., & Edmunds, R. (1999). Colour and polarity contributions to global motion perception. *Vision Research, 39,* 1813–1822.

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

Watson, A. B., & Ahumada, A. J., Jr. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A, 2,* 322–341.

Wenderoth, P. (1996). The effects of the contrast polarity of dot-pair partners on the detection of bilateral symmetry. *Perception, 25,* 757–771.

Wilson, J. A., Switkes, E., & De Valois, R. L. (2004). Glass pattern studies of local and global processing of contrast variations. *Vision Research, 44,* 2629–2641.