Binocular vision enhances phase discrimination by filtering the background

BRUCE A. SCHNEIDER, GIAMPAOLO MORAGLIA, and FILIPPO SPERANZA
University of Toronto, Mississauga, Ontario, Canada

Previous studies have shown that the detectability of a noise-masked target can be enhanced under stereoscopic viewing when the target's interocular disparity differs from that of the noise. This enhanced detectability can be accounted for by a model postulating that the binocular system linearly sums the left-eye and right-eye views of a visual scene. This model also predicts enhanced phase discrimination under specifiable interocular disparities of target and noise. Two experiments were conducted in which subjects were asked to discriminate between two luminance patterns (target and foil) that differed only in phase. The target patterns were constructed by summing two vertical sinusoidal gratings in which the phase difference between the higher and the lower spatial frequency gratings was 45° . The foils contained the same two component frequencies, with a phase difference of -45° . Thus, targets and foils were mirror images of one another. The ability of subjects to discriminate between these stereoscopically viewed mirror-image patterns was investigated under two sets of interocular disparities: those that, according to our model, would unmask one or both spatial frequency components, and those that would leave both components masked by the noise. Phase discrimination was enhanced only when both component frequencies of the target and foil were unmasked. The implications of these findings for template-matching and phase-discrimination models of pattern discrimination are considered.

It is difficult to visually detect and identify an object whose luminance pattern is similar to that of the background. Indeed, there are many examples in nature of animals whose body shape and coat have evolved in such a way as to allow them to blend easily into their natural habitat. Of course, the evolutionary development of camouflage in prey animals has been accompanied by complementary countermeasures in predators. Thus, it has been argued (e.g., Julesz, 1971) that binocular vision may have evolved in predator species to help them break the camouflage of their prey by making use of interocular depth cues. However, it is still far from clear how this feat is accomplished by the mammalian visual system.

Binocular Unmasking

In an attempt to understand how the human binocular system may be able to defeat camouflage, we have probed this system's ability to use interocular cues to enhance the detectability of a visual pattern embedded in a noisy background. Specifically, we have shown (Moraglia & Schneider, 1990, 1991, 1992; Schneider & Moraglia, 1992, 1994; Schneider, Moraglia, & Jepson, 1989; Speranza, Moraglia, & Schneider, 1995; see, also, Henning & Hertz, 1973, 1977) that the detectability of a visual sig-

nal masked by background visual noise is considerably higher when the stimulus complex contains some interocular cues that aid in the recovery of the signal than when such cues are absent. This phenomenon has been termed *binocular unmasking*, to emphasize its functional similarity to its well-known auditory counterpart, binaural unmasking (Durlach & Colburn, 1978). In our standard experimental paradigm, the left- and right-eye images of a Gaussian-enveloped sinusoidal signal of the Gabor type are stereoscopically viewed against a noise background. Such a stimulus complex is displayed under two interocular conditions. In one condition (binoptic), the interocular displacement (if any) in target location is identical to the interocular displacement in the location of the background noise. In the other condition (dichoptic), the interocular displacement of the target differs from that of the background.

Figure 1 presents an example of dichoptic stimuli employed in a two-interval, two-alternative, forced-choice procedure. The Gabor signal to be detected (peak spatial frequency $\epsilon_l = 0.97$ c/deg) is shown at the top of the figure. The middle portion of this figure shows a sample of the Gaussian noise presented to the left and right eyes during the no-signal interval. The two-dimensional Gaussian noise field in each eye is enclosed in a square frame of uniform noise whose root-mean square (RMS) contrast differs from that of the Gaussian noise field. In this case, the right noise field, relative to the surrounding noise frame, is a right-shifted version of the left noise field (shift = 30.93 min of arc). Consequently, under stereoscopic viewing, the Gaussian noise background will appear to be located behind the noise frame. The bottom

This research was supported by grants to G.M. and B.S. from the Natural Sciences and Engineering Research Council of Canada. We thank Stan Hamstra for his comments on an earlier version of the manuscript. Correspondence concerning this article should be addressed to B. A. Schneider, Department of Psychology, Erindale Campus, University of Toronto, Mississauga, ON L5L 1C6, Canada (e-mail: bschneid@credit.erin.utoronto.ca).

Figure 1. The upper panel shows a Gabor target whose peak spatial frequency is 0.97 c/deg. The middle panels show a sample of the left-eye and right-eye displays presented stereoscopically in a detection task on a noise-alone presentation. The right-eye noise is identical to the left-eye noise but is shifted by $d_{xg} = 30.93$ min within the square frame. In the bottom panels, the Gabor signal has been added to the center of the left-eye and right-eye noise patterns.

panels show another independent sample of Gaussian noise, with the Gabor signal embedded in it. As in the middle panels, the Gaussian noise in the right eye is a right-shifted version of the noise in the left eye; the degree of shift is identical to that of the middle panels. Therefore, the noise background will appear to be located behind the uniform noise frame by the same amount as in the no-signal interval. The Gabor signal, on the other hand, appears in the same location (relative to the frame) in both the left and the right eyes. Thus, this signal will be perceived as being on the same plane as that of the frame—hence, in front of the background noise. This is a dichoptic condition, because the interocular shift in the background noise (30.93 min of arc) differs from the interocular shift in the location of the target (0 min of arc).

In the corresponding binoptic trial, the displacement in the background noise on both signal and no-signal trials is identical (30.93 min of arc). However, on signal trials, the right-eye target is right-shifted by the same amount as that for the right-eye noise (30.93 min of arc). As a result, the target will appear to be on the same plane as that of the background noise (behind the square frame) and, thus, embedded in it. This is an example of a binoptic condition, because the interocular displacement of

the target is identical to the interocular displacement of the noise.

In this paradigm, the signal contrast required for threshold detection in the dichoptic condition can be as much as 1.5 log units lower than the contrast required in the binoptic condition. However, such a significant dichoptic advantage does not always occur. In the example given above, the size of the shift in the background noise was equal to half the peak wavelength of the Gabor signal (this is not a necessary condition). Under this degree of shift, a substantial dichoptic advantage is observed. However, if the background noise is now shifted by an amount equivalent to a whole wavelength of the Gabor signal (61.86 min of arc in this example), no dichoptic advantage is observed. Thus, dichoptic viewing alone is not sufficient to explain binocular unmasking, because the amount of unmasking is critically dependent on the relative difference between the displacement of the signal and the displacement of the noise.

A Summation Model for Binocular Unmasking

Schneider and Moraglia (1992, 1994) have shown that a linear summation model can account for the pattern of results obtained in these experiments, including those dichoptic conditions that do not produce binocular unmasking. According to this model, the luminance pattern of the left eye is added to the luminance pattern of the right eye. For purposes of simplification, let us assume that this summation occurs only over corresponding points on the retinas. Let $g(x, y)$ be a band-limited, two-dimensional Gaussian noise whose power spectral density function $G(\epsilon, \eta)$ is

$$G(\epsilon, \eta) = A_g, \text{ for } -\epsilon_0 \leq \epsilon \leq \epsilon_0 \text{ and } -\eta_0 \leq \eta \leq \eta_0$$

$$G(\epsilon, \eta) = 0, \text{ elsewhere,} \quad (1)$$

where x and y are the spatial coordinates of the display, A_g is the spectrum level of the noise, ϵ and η are the spatial frequency variables corresponding to the horizontal and vertical axes, and ϵ_0 and η_0 are the spatial frequency limits of the band-limited noise. Assume that a noise field is presented to the left eye, $g(x, y)$, and an identical horizontally shifted version, $g(x + d_{xg}, y)$, to the right eye, where d_{xg} corresponds to the magnitude of the horizontal displacement of the noise. A linear summation of the left and right noise fields—that is,

$$[g(x, y) + g(x + d_{xg}, y)] \quad (2)$$

—will result in a spectral density function, $F(\epsilon, \eta)$,

$$F(\epsilon, \eta) = G(\epsilon, \eta)[2 + 2\cos(2\pi\epsilon d_{xg})], \quad (3)$$

that varies sinusoidally with horizontal spatial frequency, with troughs at odd multiples of $1/(2d_{xg})$ [i.e., at spatial frequencies $1/(2d_{xg})$, $3/(2d_{xg})$, etc.] and peaks at even multiples of $1/(2d_{xg})$ [i.e., at spatial frequencies $1/d_{xg}$, $2/d_{xg}$, etc.]. To appreciate why the spectrum of $[g(x, y) + g(x + d_{xg}, y)]$ is notched, recall that a band-limited two-dimensional Gaussian noise can be represented by a sum

of sinusoidal gratings whose amplitudes, phases, and orientations are random. The vertical sinusoidal grating whose wavelength is $2d_{xg}$ will be, in the case of a displacement equal to d_{xg} , 180° out of phase in the two eyes. This sine wave will, therefore, be canceled by the linear summation of the monocular inputs. Conversely, the sine wave with wavelength d_{xg} will be in phase in the two eyes. Summation will thus produce a sinusoidal wave of the same frequency and twice the amplitude. All the sinusoidal components whose spatial frequency is neither an even nor an odd multiple of $1/(2d_{xg})$ will be either partially summed or partially canceled.

Figure 2 shows the spectral density function along the horizontal spatial frequency axis for the summated noise obtained with a horizontal displacement of $d_{xg} = -30.93$ min of arc. Note that it varies sinusoidally with troughs at 0.97, 2.91, 4.85, and 6.79 c/deg and peaks at 1.94, 3.88, 5.82, and 7.76 c/deg. Suppose now that we were to present a Gabor pattern centered within the left and right frames, so that its image fell on corresponding points on the retinas—that is, the shift of the right-eye target, d_{xt} , relative to the left-eye target, was 0. In this instance, summation of the left- and right-eye images would simply double the amplitude of the Gabor. Now, if the center frequency of the Gabor were 0.97, 2.91, 4.85, or 6.79 c/deg, the signal should be relatively easy to detect, because there would be no energy in the summated noise pattern at each of these spatial frequencies. Assume a binocular spatial frequency channel tuned to one of these frequencies—for example, 0.97 c/deg. When the noise alone is presented, this channel will be relatively quiescent, because binocular summation will result in complete can-

cellation of the 0.97 c/deg component of the noise and in partial cancellation of neighboring spatial frequencies. The activity in this channel, then, will be a function of the total amount of noise falling within its bandwidth, plus a small contribution from internal noise (baseline activity in the channel in the absence of external stimulation). The activity in this channel, however, will be markedly increased if a Gabor with peak frequency of 0.97 c/deg is added in phase to the noise in both eyes. Thus, a comparison across intervals of the activity level in this channel will indicate whether the Gabor pattern is present.

Note that in the binoptic case, the right-eye Gabor pattern will be displaced by the same amount as the right-eye noise pattern (i.e., $d_{xt} = d_{xg}$). Therefore, if the peak spatial frequency in the Gabor is at 0.97 c/deg, this spatial frequency will be canceled during the summation process. If sufficiently intense, however, this spatial frequency should be detectable by a monocular channel tuned to it (or, as we will show later, by binocular channels tuned to a specific nonzero disparity). However, since this spatial frequency will be masked by the noise in this channel, the contrast required for detection should be greater in the binoptic than in the dichoptic condition. Thus, the decibel difference between the signal-to-noise ratio required for detection in the binoptic case and the signal-to-noise ratio required for detection in the dichoptic case will provide a measure of the degree of unmasking obtained by making use of binocular cues. We refer to this difference as a *binocular masking level difference* (BMLD).

In the above example, we described what would happen if we presented a Gabor pattern whose peak spatial

Figure 2. Variation in spectral power (arbitrary units) for the sum of left-eye and right-eye Gaussian noise fields ($d_{xg} = 30.93$ min) as a function of horizontal spatial frequency.

frequency corresponded to a trough in the spectrum of the summated noise. If, on the other hand, we were to present a Gabor pattern whose peak spatial frequency corresponded to a peak in the spectrum of the summated noise, summation would simply double the amplitude of both Gabor signal and noise, leaving the signal-to-noise ratio unchanged from its value in either monocular channel. Thus, Gabor patterns with peak spatial frequencies corresponding to the peaks in the spectrum of the summated noise should be as difficult to detect under binoptic conditions as under monocular conditions.

We have shown (Schneider & Moraglia, 1994) that a generalization of this summation model can account for the detection of Gabor stimuli embedded in background noise for horizontal shifts in either noise or target (Moraglia & Schneider, 1990; Schneider et al., 1989; Speranza et al., 1995), vertical shifts in either noise or target (Moraglia & Schneider, 1991, 1992), or combinations of horizontal and vertical shifts (Schneider & Moraglia, 1994). In this generalization, we assume, in addition to a zero-disparity channel where summation occurs at corresponding retinal locations, that summation also occurs in channels with nonzero disparity. A simple illustration of this model is shown in Figure 3. The circles depict points on the two retinas. Circles that have identical reti-

nal coordinate values identify corresponding points on the two retinas. Let us assume that corresponding points are added to form the zero-disparity channel (this is the channel that we referred to specifically in the presentation of the model so far). A second, parallel, disparity channel could then be formed by adding point (x, y) in the left eye to point $(x + d_\alpha, y)$ in the right eye. For example, if $d_\alpha = -2$, the summation in this channel is equivalent to shifting the right-eye pattern 2 units to the right before adding it to the left-eye pattern. Thus, the point $(0, 0)$ in the left eye would be added to the point $(-2, 0)$ in the right eye, and so on for all the retinal points. Note that a horizontal shift to the right is indicated by a negative value of d_α . Of course, if the right-eye shift were to the left by 2 units, the value of d_α would be $+2$. For similar reasons, a vertical shift upwards in the right-eye pattern corresponds to a negative value of d_β , whereas a downwards shift produces a positive value of d_β . Other disparity channels can then be formed by adding horizontal and vertical shifts within a limited range; such channels can be assumed to be operating in parallel.

According to our model, detection is based on the output of the disparity channel that has the highest signal-to-noise ratio. For instance, in terms of the model, the highest signal-to-noise ratio for a Gabor signal in the ab-

Figure 3. Top panel: Retinal points in the left and right eyes are indicated by circles whose horizontal and vertical positions, relative to the center of the fovea, are specified by the horizontal and vertical axes. Bottom panel: For one row of the left-eye and right-eye points, the left-eye and right-eye inputs to the disparity channel defined by ($d_\alpha = -2$, $d_\beta = 0$) are indicated.

sence of noise would be in the zero-disparity channel when the Gabor is presented on corresponding points on the two retinas (and is, thus, in phase). The lowest signal-to-noise ratio would be achieved in the disparity channel whose disparity is equal to half the wavelength of the Gabor. Conversely, the highest signal-to-noise ratio would occur in this same channel when the Gabor is presented in counterphase to the two eyes. Finally, it is assumed that there is some internal noise in each of these disparity channels, so that detection cannot occur at infinitesimally small values of contrast (Schneider & Moraglia, 1992) and there is some degree of interocular jitter that will interfere somewhat with perfect summation (Speranza et al., 1995).

In the dichoptic condition of the current example ($d_{xg} = -30.93$ min of arc, $\epsilon_i = 0.97$, $\eta = 0$ c/deg), the Gabor target will have the highest signal-to-noise ratio in the zero-disparity channel. In the binoptic condition, the signal level is negligible in the zero-disparity channel because d_{xg} also equals -30.93 min of arc. However, the signal-to-noise ratio in the binoptic condition for the channel defined by ($d_\alpha = 30.93$ min of arc, $d_\beta = 0$ min of arc) will be approximately the same as that in the monocular channel, because, in this disparity channel, both signal and noise are in phase (see the Appendix, section 1). Therefore, in the binoptic condition, the amplitude of both signal and background will be doubled in this channel. Consequently, in the binoptic condition, the detection threshold should be approximately the same as that for

monocular detection, whereas in the dichoptic condition, the detection threshold should be considerably lower.

Can Binocular Unmasking Facilitate Pattern Discrimination?

Our previous studies have shown that the binocular system can exploit, under appropriate circumstances, interocular disparity differences between masker and target to enhance the latter's detectability. However, none of these studies addressed the question of whether binocular processing can also enhance the visual system's ability to discriminate among different target patterns.

Complex visual targets can be decomposed into a set of component spatial frequencies differing in amplitude, phase, and orientation. As many studies have shown (e.g., Bennett, 1993; Bennett & Banks, 1991; Lawden, 1983; Meese, 1994; Piotrowski & Campbell, 1982), the phase relationships among these spatial frequency components are critical for target discrimination and recognition. Hence, if binocular unmasking is to enhance the visual system's ability to carry out such tasks, it should enhance its ability to discriminate phase differences between the target's component frequencies. The investigation of this hypothesis motivated the present study.

In Experiment 1, the subjects were asked to discriminate between two patterns. The patterns were formed by adding two vertical sinusoidal gratings whose spatial frequencies were ϵ_i and $3\epsilon_i$, the phase relationship between the two gratings being 45° for the target stimulus and -45°

Figure 4. Left panels: Luminance profiles for the four targets in Experiment 1 are shown (top of each panel), along with the luminance profiles of their two component frequencies (0.97 c/deg at the bottom of each panel and 2.91 c/deg in the middle of each panel). Right panels: Corresponding luminance profiles for the four foils in Experiment 1, along with the luminance profiles of their component frequencies.

for the foil. Figure 4 shows the luminance profiles of the targets and foils used in these experiments, along with their constituent gratings. Note that the four targets and the four foils differ only in the locations of their peaks and that targets and foils are mirror images of one another. The right-eye noise was shifted to the left by $1/(2\epsilon_r)$ in both dichoptic and binoptic conditions. As is shown in Figure 2, this has the effect of creating notches in the summated noise spectrum at both ϵ_l and $3\epsilon_l$ in the zero-disparity channel. Therefore, both frequencies should be unmasked. The target and foil patterns in the dichoptic condition were centered in the left- and right-eye frames so that their component frequencies would sum in the zero-disparity channel. Therefore, the component frequencies making up the target and foil patterns should be more easily detected in the dichoptic condition than in the binoptic condition, where the right-eye image of the target is right-shifted by the same amount as the noise (see the Appendix, section 1). The question addressed in this experiment is whether the unmasking of these two component frequencies will result in enhanced pattern discrimination.

In Experiment 2, the lateral displacement of the noise was the same as that in Experiment 1. The two patterns, however, were mirror images based on ϵ_l and $2\epsilon_l$ (see Figure 5). Because the displacement of the noise was the same as that in Experiment 1, the ϵ_l grating should be unmasked, but not the $2\epsilon_l$ grating. The question here is whether unmasking only one of the spatial frequency components

of the pattern will enhance pattern discrimination. Both phase discrimination and template-matching models predict a dichoptic advantage even when one only of the spatial frequencies is unmasked (see the Appendix, parts 2 and 3).

METHOD

Subjects

Two males (G.M., one of the authors, and S.A.) and two females (J.C. and C.L.) served as observers in both experiments. All the observers had normal or corrected-to-normal acuity and normal stereo vision. G.M. and J.C. were experienced psychophysical observers, and all but G.M. were unaware of the hypothesis under investigation.

Apparatus and Stimuli

The stimuli were generated by a microcomputer equipped with an ITI FG-100 image processor with 512×512 pixel spatial resolution (8-bit gray scale level) and were presented on a BW 9-in. TV monitor (Panasonic WV 5370). Displays were seen through a simple lens stereoscope. A rectangular aperture in an opaque-black metallic frame limited the visible portion of the monitor to a rectangular window. A headrest was used to constrain head movements and to help maintain fixation. The subjects viewed the left- and right-eye fields through neutral density filters. Under these viewing conditions, the distance between two adjacent pixels subtended 3.44 min of visual angle.

The left-eye display consisted of a 140×140 pixel field of two-dimensional Gaussian noise surrounded by a 15-pixel-wide square frame of uniform noise (bandwidth = 8.875 c/deg in both cases). The right eye saw a horizontally right-shifted version of the same

Figure 5. Left panels: Luminance profiles for the four targets in Experiment 2 are shown (top of each panel), along with the luminance profiles of their two component frequencies (0.97 c/deg at the bottom of each panel and 1.94 c/deg in the middle of each panel). Right panels: Corresponding luminance profiles for the four foils in Experiment 2, along with the luminance profiles of their component frequencies.

two-dimensional Gaussian noise. In all cases, the amount of the shift was $d_{xg} = -30.93$ min of arc. A gamma-corrected look-up table was used to ensure a linear relation between gray levels and luminance, which ranged from 0.1 cd/m² (gray level = 0) to 8.4 cd/m². The uniform noise was obtained by adding random deviates from -40 to +40 to the mean gray level (gray level = 128). Its RMS contrast, defined as the RMS luminance divided by the mean luminance, was 0.18. To produce the Gaussian noise, the gray values of a 140×140 matrix of random normal deviates were rescaled so that their standard deviation was equal to 80 and were peak-clipped to ensure that they ranged from -127 to +127; to the values thus obtained, the value of 128 was added. After peak-clipping, the RMS contrast of the Gaussian noise was 0.626.

Targets and foils were presented within a 72×72 pixel window that was added to the background noise in both the left- and the right-eye display. The deviation from mean gray level for the target was defined as

$$A \cdot \cos[2\pi\epsilon_i(x + d_k)] + B \cdot \cos\left[2\pi n\epsilon_i(x + d_k) + \frac{\pi}{4}\right]. \quad (4)$$

The deviation from mean gray level for the foil was defined as

$$A \cdot \cos[2\pi\epsilon_i(x + d_k)] + B \cdot \cos\left[2\pi n\epsilon_i(x + d_k) - \frac{\pi}{4}\right]. \quad (5)$$

Note that the two patterns differ only with respect to the phase relationship between ϵ_i and $n\epsilon_i$. As was stated, the value of n was 3 in Experiment 1 and 2 in Experiment 2. For both target and foil the value of A was set to a gray level that produced a contrast of 10% in both experiments. In Experiment 1, four different values of d_k were employed $[0, 1/(4\epsilon_i), 1/(2\epsilon_i), \text{ and } 3/(4\epsilon_i)]$ to produce four exemplars of target and background. Thus, the other three exemplars of the target and foil look like shifted versions of those shown in Figure 4. Note that changing d_k simply shifts the locus of the pattern, without affecting the phase relationship between ϵ_i and $3\epsilon_i$.

These four values of d_k were chosen to ensure that the two patterns could not be discriminated on the basis of their individual frequency components. If only one exemplar were employed (say, $d_k = 0$), the two patterns could, in principle, be distinguishable from one another simply on the basis of the absolute phase of the $3\epsilon_i$ grating. For example, the subjects could attend to, say, the left edge of the $3\epsilon_i$ grating and base their judgment on whether the left edge of the pattern was bright or dark (see Figure 4). Setting d_k to these values produced four distinct pairs of phases for the ϵ_i and $3\epsilon_i$ gratings for target and foil. For the targets, the paired phases [ϵ_i phase first, $3\epsilon_i$ phase second] were $[0, \pi/4]$, $[\pi/2, (7\pi)/4]$, $[\pi, (5\pi)/4]$, and $[(3\pi)/2, (3\pi)/4]$. For the foils, the paired phases were $[0, (7\pi)/4]$, $[\pi/2, (5\pi)/4]$, $[\pi, (3\pi)/4]$, and $[(3\pi)/2, \pi/4]$. Note that the two patterns cannot be distinguished on the basis of ϵ_i alone, because the four exemplars of the target are identical to the four exemplars of the foil. The same also holds true for $3\epsilon_i$. Thus, the only basis for discriminating target from foil is the phase relationship between ϵ_i and $3\epsilon_i$.

Experiment 2 (ϵ_i and $2\epsilon_i$) also used four exemplars of both target and foil. These exemplars were produced by employing four different values of d_k $[0, 1/(8\epsilon_i), 1/(4\epsilon_i), 3/(8\epsilon_i)]$. This resulted in four targets whose paired phases (ϵ_i phase, $2\epsilon_i$ phase) were $[0, \pi/4]$, $[\pi/4, (3\pi)/4]$, $[\pi/2, (5\pi)/4]$, and $[(3\pi)/4, (7\pi)/4]$. The paired phases for the foils were $[0, (7\pi)/4]$, $[\pi/4, \pi/4]$, $[\pi/2, (3\pi)/4]$, and $[(3\pi)/4, (5\pi)/4]$. As in Experiment 1, targets cannot be discriminated from foils on either ϵ_i alone or $2\epsilon_i$ alone but must be based on the phase difference between the two.

The contrast of the higher frequency component was varied in order to determine the threshold for discriminating between the two patterns under binoptic and dichoptic viewing conditions. In the binoptic condition, $d_{xt} = d_{xg}$; that is, the right-eye target was shifted by the same amount as the right-eye noise. In the dichoptic condition, $d_{xt} = 0$; that is, the right-eye target was presented in the same posi-

tion, relative to the frame, as the left-eye target. The wavelength of the lower frequency component in both experiments was 18 pixels, which corresponds to a spatial frequency of 0.97 c/deg. Each signal was presented at four contrast levels selected individually for each observer in a preexperimental session.

Procedure

A temporal two-alternative forced-choice procedure was used to estimate the contrast of the higher spatial-frequency component of the pattern that yielded correct discrimination between target and foil on 75% of the trials. Prior to the initiation of each trial, the subjects saw a display containing a 5×5 pixel fixation point in its center to aid fusion. This pretrial display also showed a sample of the target above the fixation point and a sample of the foil below it. No background noise was presented during the pretrial display, and targets and foils were presented at contrast levels identical to those to be presented on the subsequent trial. However, the positions of the pretrial target and foil were independent of the positions of the target and foil on the subsequent trial. Thus, the subjects were given information concerning target and foil contrast but not target and foil position prior to each trial.

Upon termination of the free inspection of the above display, the subjects initiated a trial by pressing a button on a response console. This caused, after 500 msec, two other displays, both containing the Gaussian noise field surrounded by uniform noise frames, to be presented in temporal succession. The first display lasted for 1 sec and was followed, after 500 msec, by another 1-sec display. The Gaussian noise fields presented in the two intervals were randomly chosen from a set of 36 such fields and differed in the two intervals and across trials. When the target was added to the noise field in the first interval, the foil was presented in the second, and vice versa. The subjects responded by pressing a button corresponding to the interval they thought contained the target. Feedback was provided for correct answers by means of a computer-generated tone. Each subject completed 768 trials in each experiment, divided into eight blocks; 96 trials were collected at each level of signal contrast (4) in both binoptic and dichoptic conditions. In each block of trials, all the displays belonged to the same condition. Before the onset of the experiment proper, each subject was given the opportunity to practice the experimental task in the various conditions to be tested. These practice sessions, which varied somewhat in both number and duration from subject to subject, were also used to select the range of contrast values to be used in the experiment.

RESULTS

Figure 6 shows that the percentage of times the target pattern was correctly identified in the two-interval forced-choice task increased as a function of the contrast of the higher frequency component for all 4 subjects in the binoptic conditions of Experiments 1 and 2. Figure 6 also indicates that it was more difficult to discriminate between the two patterns when the higher frequency component was $3\epsilon_i$ (Experiment 1), rather than $2\epsilon_i$ (Experiment 2). The average contrast of the higher harmonic at threshold (75% correct identification) was 13.5% in Experiment 1 and 8.4% in Experiment 2.

Figure 7 plots the percentage of correct responses as a function of the contrast of the $3\epsilon_i$ component for both dichoptic and binoptic conditions in Experiment 1. For all 4 subjects, percent correct rises more rapidly with increases in contrast under dichoptic than under binoptic conditions. The average contrast required for threshold

Figure 6. Percent correct as a function of signal contrast for the binoptic conditions in Experiments 1 and 2. In both experiments, the lower frequency grating had a spatial frequency of 0.97 c/deg. In Experiment 1 (squares), the spatial frequency of the higher frequency component was 2.91 c/deg; in Experiment 2 (circles), the spatial frequency of the higher frequency component was 1.94 c/deg. The data are plotted separately for each of the 4 subjects.

Figure 7. Percent correct as a function of signal contrast for the binoptic (filled squares) and dichoptic (unfilled squares) conditions in Experiment 1. The spatial frequencies of the two component gratings in both conditions were 0.97 and 2.91 c/deg. The data are plotted separately for each subject.

discrimination in the binoptic condition was 13.5%; it was 5.6% in the dichoptic condition. Thus, the contrast required for pattern discrimination under dichoptic conditions is only 40% of that required under binoptic conditions.

Figure 8 plots the percentage of correct responses as a function of the contrast of the $2\epsilon_i$ component for both dichoptic and binoptic conditions in Experiment 2. Figure 8 indicates that there is very little difference between the

Figure 8. Percent correct as a function of signal contrast for the binoptic (filled circles) and dichoptic (unfilled circles) conditions in Experiment 2. The spatial frequencies of the two component gratings in both conditions were 0.97 and 1.94 c/deg. The data are plotted separately for each subject.

binoptic and the dichoptic conditions. The average contrast required for pattern discrimination is 8.4% and 7.1% for binoptic and dichoptic presentations, respectively.

DISCUSSION

Binocular Processing and Pattern Discrimination

In Experiment 1, the subjects were asked to discriminate between two phase relations in a pattern composed of two sinusoidal gratings whose spatial frequencies were 0.97 and 3×0.97 c/deg under two conditions. In the binoptic condition, the interocular disparity in the target was identical to the interocular disparity of the noise within which it was embedded. Under such conditions, our model leads us to expect no binocular advantage. Thus, the binoptic condition represents a baseline against which we can measure the effectiveness of binocular unmasking for pattern discrimination. In the dichoptic condition, since the interocular disparity in the noise was equal to half of the wavelength of the two patterns' lower spatial frequency, this should have the effect of canceling out the noise at ϵ_1 and $3\epsilon_1$. Thus, these two frequency components of the pattern employed in Experiment 1 should have been both unmasked. If, in this task, pattern discrimination is carried out after binocular fusion, we would expect to find lower thresholds for pattern discrimination in the dichoptic than in the binoptic condition. As Figure 7 shows, that is what we found.

In Experiment 2, the subjects were asked to discriminate between a pattern composed of two vertical sinu-

soidal gratings, whose component frequencies stood in a ratio of 2:1, and the mirror image of this pattern. In the dichoptic condition, the disparity in the noise was such that it should have unmasked the lower frequency in the pattern, but not its higher frequency component. Of course, in the binoptic condition, neither frequency should have been unmasked. Because efficient discrimination requires the processing of the phase difference between the two component frequencies (one of which is masked in the dichoptic condition), we would not expect the dichoptic condition to differ as much from the binoptic condition as it did in Experiment 1. With the possible exception of subject J.C., there is no indication of a dichoptic advantage in Experiment 2. This finding is interesting also in view of the fact that both feature detection models and phase discrimination models predict some advantage in these conditions (see below).

What are the implications of these results for camouflage breaking by the binocular system? Our previous work has clearly shown that differences in interocular disparity between target and background can significantly enhance detectability. However, our present results suggest that, in order for such differences in interocular disparity to aid in pattern discrimination and pattern recognition, all (or most) of the spatial frequencies that are critical for pattern recognition must be unmasked. Hence, because differences in interocular disparity can unmask some, but not all, of the critical spatial frequencies, it is likely that stimulus situations will arise in which insufficient spatial frequency information is

available for target recognition, even though there is ample information for target detection. In situations such as these, the observer might easily detect a target, while being unable to identify it.

The present findings also have implications for the level at which pattern recognition occurs. Consider, for example, subject J.C. when she was presented, in Experiment 1, with the task of discriminating between a pattern (composed of sinusoidal gratings whose spatial frequencies were ϵ_i and $3\epsilon_i$; see Figure 4) and its mirror image when the contrast of the higher spatial frequency was set at 7.5%. In the binoptic case, which we have argued is similar to monocular viewing, her probability of successfully discriminating between the two patterns was about 60%. However, in the dichoptic condition, in which both spatial frequencies were unmasked, discrimination rose to 90% correct. Clearly, binocular unmasking facilitated this discrimination. Because this discrimination depended solely on the phase relationship between the two frequencies, it seems clear that pattern discrimination in Experiment 1 took place subsequent to the binocular integration of the images from the two eyes.

Phase Encoding, Local Contrast, Spatial Features, and Pattern Matching

To what extent can the results of these experiments be predicted by models that postulate that similar tasks are carried out by analyzing a set of local spatial cues (e.g., Babcock, 1984a, 1984b; Georgeson, 1990; Hess & Pointer, 1987; Meese, 1994; Watt, 1985; Yuille & Poggio, 1988)? An examination of Figures 4 and 5 shows that it would be hard to explain discrimination in these experiments in terms of an analysis of local contrasts between adjacent maxima and minima of luminance, because the only difference between target and foil was that one was the exact mirror image of the other. Thus, the absolute magnitude of spatial contrast was always the same for both target and foil.

Of course, discrimination could be based on the direction of contrast change. Figure 4 shows that the trough to the left of a peak in the target's luminance profile is more pronounced than that to the right of the peak. Consequently, the patterns could be discriminated on the basis of whether the peak-to-trough contrast was greater, with respect to the first trough, to the left or the right of a peak. Note, however, that in order for such a strategy to work, the location of the peak must be known first. This was difficult to do in our experiments, because the location of a peak varied between target and foil and from target to target and depended on the degree of contrast in the higher frequency component of the compound grating. Moreover, once random noise is added to the pattern, it is not clear how the location of a peak is to be specified. If the observer, for example, focused on a secondary peak of the target, the peak-to-nearest-trough contrast would be greater on the right than on the left. Thus, it is difficult to see how a local contrast analysis would be able to account for this discrimination.

It is also difficult to see how vertical zero crossings (Meese, 1994; Yuille & Poggio, 1988) could be used to facilitate the discrimination. The amount of random noise added to each line of the horizontal display is such that there is a large number of zero crossings for each such line. It is, of course, possible that phase discrimination might be related to the relative location of zero crossings after spatial frequency filtering. However, without specifying the nature of such filtering, it is difficult to construct a model of phase discrimination on the basis of the information contained in the zero crossings.

It is possible that more global features in the spatial array form the basis for the discrimination. The most global feature analyzer that could be employed would be one in which the analyzer exactly matched the luminance profile of the target. Such a template-matching model (see, e.g., Burgess & Ghaudeharian, 1984; Caelli & Moraglia, 1985, 1986) includes all the local features. To evaluate the performance of such a pattern-matching device (see the Appendix, section 2, for a mathematical formulation of the model), we assumed an observer who had complete knowledge of the luminance profiles of both target and foil but did not know in which of four positions the target and/or the foil would occur. Consequently, the observer cross-correlated each of eight different patterns (four for the targets, and four for the foils) with the stimulus display in each of the two intervals of the two-interval forced-choice procedure. This observer labeled the pattern in the first interval a target, if the highest cross-correlation product came from one of the target patterns, and a foil, if the highest product came from one of the foil patterns. A similar labeling procedure was carried out for the second interval. If this cross-correlation procedure resulted in one of the interval's being identified as containing a target while the other contained a foil, the observer identified the interval labeled target as containing the target. If cross-correlation indicated that both intervals contained a target, the observer chose the interval producing the highest cross-correlation product as the one most likely to contain the target. If cross-correlation indicated that both intervals contained a foil, the observer chose the interval with the smallest cross-correlation product as the one most likely to contain the target. Finally, it was assumed that this observer made use of the entire display in performing the cross-correlation. This is akin to assuming that the effective field of view for pattern discrimination is at least 72×72 pixels, or 4.13×4.13 squared degrees.

The template model can be applied to both binoptic and dichoptic conditions. In the binoptic condition, the templates are cross-correlated with the left-eye display. (Since the interocular relationships between pattern and masker are the same in both the left and the right eye, it does not matter which one is used.) In the dichoptic case, the templates are cross-correlated with the sum of the left- and right-eye displays to permit binocular unmasking (see the Appendix, section 2, for details). We refer to this model as the *unconstrained* template-matching ob-

Figure 9. Simulated performance for the unconstrained template-matching and Bennett observers. Upper left quadrant: performance of the Bennett observer in dichoptic and binoptic conditions for $\epsilon + 2\epsilon$. Upper right quadrant: performance of the template-matching observer in dichoptic and binoptic conditions for $\epsilon + 2\epsilon$. Lower left quadrant: performance of the Bennett observer for $\epsilon + 3\epsilon$. Lower right quadrant: performance of the template-matching observer for $\epsilon + 3\epsilon$. In the lower quadrants, only binoptic performance is shown, since dichoptic performance is perfect (see the text).

server, to differentiate it from the model of an observer who follows the same detection strategy but is constrained by a number of limitations, as will be described below.

Figure 9 (upper right quadrant) plots psychometric functions for this pattern-matching device under binoptic and dichoptic conditions for the ϵ and 2ϵ phase discrimination. The template-matching device not only is significantly better than a human observer, but it also shows a dichoptic advantage. As is shown in the Appendix (section 2), this advantage arises because the lower frequency component is completely unmasked by binocular summation.

The lower right quadrant of Figure 9 displays the psychometric function obtained from the same template-matching device for the ϵ and 3ϵ case under binoptic presentation. The dichoptic psychometric function is not shown, since performance would be perfect for any non-zero contrast at 3ϵ . Hence, when both frequency components are unmasked, the performance of the template-matching device is error free.

Bennett (1993) proposed a model in which phase-reversal discrimination for compound gratings is based on the output of sine and cosine mechanisms (i.e., on the output of phase detectors). In phase-reversal experiments of this sort (Bennett, 1993; Bennett & Banks, 1991; Field & Nachmias, 1984; Meese, 1994; Morrone, Burr, & Spinelli, 1989), observers are asked to discriminate between a two-component compound grating (spatial frequencies, ϵ and $n\epsilon$, where n is an integer) and the same compound grating in which the phase of the $n\epsilon$ component is reversed. According to Bennett's model, the ob-

server first locates peaks in the luminance profile for the lower spatial frequency and then centers sine and cosine filters tuned to a spatial frequency of $n\epsilon$ cycles/deg at these peaks. Discrimination is based on the difference in the outputs of these sine and cosine filters for the original pattern and for the phase-reversal pattern. Bennett does not specify how the observer centers these filters at the locus of a peak in the lower spatial frequency. Instead, he simply assumes that there is some degree of error in locating such a peak. Whereas this may not be a problem for an unmasked pattern, a way has to be found to determine the locations of peaks in the luminance pattern of the lower spatial frequency when there is a considerable amount of noise present.

We have therefore developed a model in which the observer uses sine and cosine filters tuned to the lower spatial frequency in order to identify the loci of peaks at this frequency (see the Appendix, section 3, for further details). This observer then places a sine filter tuned to the higher spatial frequency at the estimated locus of a peak in the lower spatial frequency. To discriminate between targets and foils, the observer simply determines in which of the two intervals of the forced-choice procedure the signed output of the sine filter is lowest. (In this case, phase discrimination is based on the output of the sine filter only, because the cosine filter's output does not differentiate between a $+45^\circ$ phase shift and a -45° phase shift.) This modified Bennett model differs from Bennett's (1993) original model in two ways. First, the mechanism for identifying a local peak in the lower frequency grating is specified. Second, there is no internal noise in

the model. (The external noise is so much greater than any presumed internal noise that the latter is unlikely to affect the process. Of course, if internal noise is added to the model, it will be able to predict phase discrimination in the absence of external noise.) The upper left quadrant of Figure 9 plots psychometric functions obtained from this modified and unconstrained Bennett observer for binoptic and dichoptic conditions for the ϵ and 2ϵ pattern discrimination; the lower right quadrant of Figure 9 plots the psychometric function for the binoptic case for the ϵ and 3ϵ pattern discrimination. Note that the performance of the unconstrained Bennett observer is similar to that of the corresponding template-matching device and that the performance of both such observers is perfect under dichoptic viewing conditions for ϵ and 3ϵ pattern discrimination.

A real observer, of course, would differ from the unconstrained observers described above in at least four important ways. First, unlike the unconstrained observers, the spectral resolution of real observers is not perfect, being limited by the bandpass characteristics of their spatial frequency filters. Second, the eyes of real observers, even under conditions of strict fixation, are always in motion. Therefore, the presence of a measure of eye jitter will prevent the process of binocular summation from being carried out faultlessly by the human observer, as is assumed to occur in the unconstrained device. Third, the unconstrained observer is not affected by internal noise, but the latter is always present in the spatial-frequency channels of real observers. However, in the present experiments, owing to the fact that the external noise is so much greater than the internal noise, it is

likely that the contribution of this factor to performance is negligible. Finally, the effective field of view for the unconstrained observer is assumed to be infinite. The effective field of view for efficient pattern discrimination in real observers is likely to be considerably smaller than that, being predominantly limited to the fovea region.

Figure 10 plots psychometric functions for the template-matching observer and for the Bennett observer when it is assumed (1) that the patterns are first processed by spatial-frequency channels whose bandwidths (between 1/2 power points) are 0.5 octaves wide; (2) that the standard deviation of eye jitter is approximately 2.1 min of arc; and (3) that the effective field of view (the field over which information can be processed) is approximately $2.06^\circ \times 2.06^\circ$ of visual angle. (Note that a bandwidth of 0.5 octaves and a standard deviation of 2.1 for eye jitter represent the lower limits of the bandwidth estimates for spatial-frequency channels and eye jitter, respectively; see, e.g., Graham, 1989; St. Cyr & Fender, 1969.) Several features of this figure should be noted. First, a comparison of Figures 6–8 with Figure 10 shows that the performance of both template-matching and Bennett observers is better than that of our 4 subjects when lower estimates of channel bandwidth and eye jitter are used. Second, under binoptic viewing, the template-matching observer achieves better discrimination performance for ϵ and 3ϵ than for ϵ and 2ϵ . (In the Appendix, we show that this advantage results from the fact that the bandwidths of spatial-frequency channels increase with frequency.) On the other hand, the performance of real observers generally worsens when the frequency of the upper component is

Figure 10. Performance of band-limited and eye-jittered template-matching and Bennett observers. Both observers were assumed to have spatial-frequency channel bandwidths of 0.5 octaves, and eye jitter whose standard deviation was 2.1 min of arc. Upper left quadrant: performance of the Bennett observer in dichoptic and binoptic conditions for $\epsilon + 2\epsilon$. Upper right quadrant: performance of the template-matching observer in dichoptic and binoptic conditions for $\epsilon + 2\epsilon$. Lower left quadrant: performance of the Bennett observer in dichoptic and binoptic conditions for $\epsilon + 3\epsilon$. Lower right quadrant: performance of the template-matching observer in dichoptic and binoptic conditions for $\epsilon + 3\epsilon$.

raised. It thus appears that, unless there are other factors that limit performance as the frequency of the upper component increases, it is unlikely that our real observers employed a template-matching strategy.

Figure 10 also shows that, in the binoptic condition, the performance of the Bennett observer is poorer for ϵ and 3ϵ than it is for ϵ and 2ϵ . This is to be expected, given that errors in locating a peak in the ϵ grating will have an increasingly deleterious effect on the discriminative capacity of the $n\epsilon$ filters, as the value of n increases. Thus, the Bennett observer shares with our real observers a decline in performance as the frequency of the upper component is increased.

Finally, Figure 10 indicates that the Bennett and template-matching models both predict a substantial advantage for the dichoptic condition. Human observers, however, achieved a dichoptic advantage only when both frequency components were unmasked.

Given that the template-matching model predicts improved performance as the frequency of the upper component is increased, whereas the data from our observers show that performance worsens as the frequency of the upper component is increased, we did not attempt to obtain a better fit of this model to the average data. We did, however, attempt to find bandwidth parameters and eye-jitter parameters that would provide a better fit to the average data of our four observers for the Bennett model. Figure 11 plots the predictions of the Bennett model for the case in which the bandwidths of the spatial-frequency channels were 1 octave in width, the standard deviation of eye jitter was 4 min of arc, and the effective field of view was again limited to the fovea ($2.07^\circ \times 2.07^\circ$). Also shown is the average performance of our 4 observers. (Note that, in the $\epsilon + 3\epsilon$ case, the 4 observers were tested at different values of target contrast. However, there were three values in common to 3 of the 4 observers. Because the psycho-

metric function of the 4th observer bracketed these values, a linear interpolation was used to estimate the performance of the 4th observer at these three values. Thus, Figure 11 plots percent correct for the average of these three contrasts for the $\epsilon + 3\epsilon$ case.) Figure 11 shows, for the binoptic condition in $\epsilon + 2\epsilon$, and for both the binoptic and dichoptic conditions at $\epsilon + 3\epsilon$, that these new parameter values produce predicted psychometric functions within the range observed in these experiments. The Bennett model, however, predicts a clear dichoptic advantage at $\epsilon + 2\epsilon$, whereas our 4 observers (with one possible exception) show no such advantage. Note that the parameter values used in the simulation for Figure 11 are well within the range found for human observers. Thus, with the exception of the lack of a clear dichoptic advantage for $\epsilon + 2\epsilon$ phase discrimination, the performance of a Bennett observer subjected to the constraints we used in our simulation roughly matches that of human observers.

The modified Bennett model would become capable of accounting for the lack of dichoptic advantage in the $\epsilon + 2\epsilon$ case if it was assumed that optimal pattern discrimination is possible only when the component frequencies are approximately equally detectable. Note that, ignoring minor differences in contrast detection thresholds for ϵ and 3ϵ , these two components had approximately equal detectability in the dichoptic condition of Experiment 1 because of noise reduction at both of these frequencies. However, in the dichoptic condition of Experiment 2, only ϵ was unmasked. Therefore, since ϵ and 2ϵ were not equally detectable, relative phase information could not be extracted easily. However, the reasons for such a constraint are not clear to us. Whether a modified Bennett model should be claimed to mediate binocular, phase-based pattern discrimination, or whether an entirely new model of this process should be developed instead, constitutes a question we propose to pursue in forthcoming studies.

Figure 11. Performance of a band-limited and eye-jittered Bennett observer in dichoptic and binoptic conditions for $\epsilon + 2\epsilon$ (upper panel) and $\epsilon + 3\epsilon$ (lower panel). The observer was assumed to have spatial-frequency channel bandwidths of 1 octave and eye jitters whose standard deviation was 4 min of arc.

REFERENCES

- BABCOCK, D. R. (1984a). How do we discriminate spatial phase? *Vision Research*, **24**, 1847-1857.
- BABCOCK, D. R. (1984b). Spatial phase or luminance profile discrimination? *Vision Research*, **24**, 613-623.
- BENNETT, P. J. (1993). The harmonic bandwidth of phase-reversal discrimination. *Perception & Psychophysics*, **53**, 292-304.
- BENNETT, P. J., & BANKS, M. S. (1991). The effects of contrast, spatial scale, and orientation on foveal and peripheral phase discrimination. *Vision Research*, **31**, 1759-1786.
- BURGESS, A., & GHAUDEHARIAN, H. (1984). Visual signal detection. I. Ability to use phase information. *Journal of the Optical Society of America*, **1**, 900-905.
- CAELLI, T., & MORAGLIA, G. (1985). On the detection of Gabor signals and discrimination of Gabor textures. *Vision Research*, **25**, 671-684.
- CAELLI, T., & MORAGLIA, G. (1986). On the detection of signals embedded in natural scenes. *Perception & Psychophysics*, **39**, 87-95.
- DURLACH, N. I., & COLBURN, H. S. (1978). Binaural phenomena. In E. C. Carterette & M. A. Friedman (Eds.), *Handbook of perception* (Vol. 4, pp. 365-466). New York: Academic Press.
- FIELD, D. J., & NACHMIAS, J. (1984). Phase reversal discrimination. *Vision Research*, **24**, 333-346.
- GEORGESON, M. A. (1990). Human vision combines oriented filters to compute edges. *Perception*, **19**, 354.
- GRAHAM, N. (1989). *Visual pattern analyzers*. New York: Oxford University Press.
- HENNING, G. B., & HERTZ, G. B. (1973). Binocular masking level differences in sinusoidal grating detection. *Vision Research*, **13**, 2455-2463.
- HENNING, G. B., & HERTZ, G. B. (1977). The influence of bandwidth and temporal properties of spatial noise on binocular masking-level differences. *Vision Research*, **17**, 399-402.
- HESS, R. F., & POINTER, J. S. (1987). Evidence for spatially local computations underlying discrimination of periodic patterns in fovea and periphery. *Vision Research*, **27**, 1343-1360.
- JULESZ, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.
- LAWDEN, M. C. (1983). An investigation of the ability of the human visual system to encode spatial phase relationships. *Vision Research*, **23**, 1451-1464.
- MEESE, T. S. (1994). Phase-reversal discrimination in one and two dimensions: Performance is limited by spatial repetition, not spatial frequency content. *Vision Research*, **35**, 2157-2167.
- MORAGLIA, G., & SCHNEIDER, B. (1990). Effects of direction and magnitude of horizontal disparity on binocular unmasking. *Perception*, **19**, 581-593.
- MORAGLIA, G., & SCHNEIDER, B. (1991). Binocular unmasking with vertical disparity. *Canadian Journal of Psychology*, **45**, 353-366.
- MORAGLIA, G., & SCHNEIDER, B. (1992). On binocular unmasking of signals in noise: Further tests of the summation hypothesis. *Vision Research*, **32**, 375-385.
- MORRONE, M. C., BURR, D. C., & SPINELLI, D. (1989). Discrimination of spatial phase in central and peripheral vision. *Vision Research*, **29**, 433-446.
- POTROWSKI, L. N., & CAMPBELL, F. W. (1982). A demonstration of the visual importance and flexibility of spatial-frequency amplitude and phase. *Perception*, **11**, 337-346.
- SCHNEIDER, B., & MORAGLIA, G. (1992). Binocular unmasking with unequal interocular contrast: The case for multiple Cyclopean eyes. *Perception & Psychophysics*, **52**, 639-660.
- SCHNEIDER, B., & MORAGLIA, G. (1994). Binocular vision enhances target detection by filtering the background. *Perception*, **23**, 1267-1286.
- SCHNEIDER, B., MORAGLIA, G., & JEPSON, A. (1989). Binocular unmasking: An analog to binaural unmasking? *Science*, **243**, 1479-1481.
- SPERANZA, F., MORAGLIA, G., & SCHNEIDER, B. (1995). Age-related changes in binocular vision: Detection of noise-masked targets in young and old observers. *Journal of Gerontology: Psychological Sciences*, **50B**, 114-123.
- ST. CYR, G. J., & FENDER, D. H. (1969). The interplay of drifts and flicks in binocular fixation. *Vision Research*, **9**, 245-265.
- STEINMAN, R. M. (1965). Effects of target size, luminance, and colour on monocular fixation. *Journal of the Optical Society of America*, **55**, 1158-1165.
- WATT, R. J. (1985). Structured representation in low-level vision. *Nature*, **313**, 266-267.
- YUILLE, A. L., & POGGIO, T. (1988). Scaling and fingerprint theorems for zero-crossings. In C. Brown (Ed.), *Advances in computer vision* (Vol. 2, pp. 47-78). Hillsdale, NJ: Erlbaum.

APPENDIX

1. A Linear Summation Model of Binocular Unmasking

Let $g(x, y)$ be the right-eye noise pattern. The left-eye noise pattern $g(x + d_{xg}, y + d_{yg})$ is assumed to be a horizontally and vertically shifted version of the right-eye noise pattern, with d_{xg} specifying the degree of horizontal shift and d_{yg} the degree of vertical shift. In addition, we assume that summation occurs over different disparity channels, with the degree of internal horizontal shift, d_α , and the degree of internal vertical shift, d_β , specifying the disparity channel (d_α, d_β). The summated left and right noise patterns in disparity channel (d_α, d_β), therefore, is

$$g(x, y) + g(x + d_{xg} + d_\alpha, y + d_{yg} + d_\beta). \quad (\text{A1.1})$$

The spectral density function for this summated pattern is

$$F[\epsilon, \eta] = G[\epsilon, \eta] \cdot (2 + 2 \cos[2\pi\epsilon(d_{xg} + d_\alpha) + 2\pi\eta(d_{yg} + d_\beta)]). \quad (\text{A1.2})$$

According to the model, the amount of the noise energy at the output of the spatial frequency filter centered at ϵ_f, η_f is affected by the bandwidth of the filter, the degree of spatial jitter in the eyes, and the level of internal noise in the channel in the absence of input.

We assume a Gabor filter whose spectral profile is

$$B[\epsilon, \eta] = \frac{1}{2} e^{-b^2(\epsilon - \epsilon_f)^2} e^{-b^2(\eta - \eta_f)^2} + \frac{1}{2} e^{-b^2(\epsilon + \epsilon_f)^2} e^{-b^2(\eta + \eta_f)^2}. \quad (\text{A1.3})$$

Therefore, the noise energy passing through the filter is given by

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F[\epsilon, \eta] B[\epsilon, \eta] d\epsilon d\eta = \frac{2\pi A_g}{b^2} \left(1 + e^{-\frac{\pi^2(d_{xg}+d_\alpha)^2}{b^2}} e^{-\frac{\pi^2(d_{yg}+d_\beta)^2}{b^2}} \cos[2\pi\epsilon_f(d_{xg}+d_\alpha) + 2\pi\eta_f(d_{yg}+d_\beta)] \right). \quad (A1.4)$$

Because there are small involuntary eye movements, even under fixation, the relative positions of the left- and right-eye noises vary randomly over time. We refer to the horizontal component of this movement as horizontal spatial jitter. We assume that this jitter is normally distributed with mean value equal to zero and a standard deviation equal to σ . Values of σ may vary between about 2 and 4 min of arc, depending on the nature of the fixation target (St. Cyr & Fender, 1969; Steinman, 1965). Therefore, the energy from the external noise passing through the filter at any one moment becomes

$$\frac{2\pi A_g}{b^2} \left(1 + e^{-\frac{\pi^2(d_{xg}+r+d_\alpha)^2}{b^2}} e^{-\frac{\pi^2(d_{yg}+d_\beta)^2}{b^2}} \cos[2\pi\epsilon_f(d_{xg}+r+d_\alpha) + 2\pi\eta_f(d_{yg}+d_\beta)] \right), \quad (A1.5)$$

where r is the extent in degrees of the momentary displacement from eye jitter. If we assume that r is normally distributed, with mean 0, and standard deviation is σ , it is easy to show that the average energy passing through the filter is

$$\frac{2\pi A_g}{b^2} \left(1 + \frac{b}{c} e^{-\frac{\pi^2 X^2}{c^2}} e^{-\frac{\pi^2 Y^2}{c^2}} e^{-\frac{2b^2\epsilon_f^2\pi^2\sigma^2}{c^2}} \cos \left[\frac{2\pi\epsilon_f b^2 X}{c^2} + 2\pi\eta_f Y \right] \right), \quad (A1.6)$$

where

$$c = \sqrt{b^2 + 2\pi^2\sigma^2}, \quad X = d_{xg} + d_\alpha, \quad Y = d_{yg} + d_\beta.$$

The model also assumes that there is uncorrelated internal noise in both the left- and the right-eye channels. If we assume that the spectral distribution of the uncorrelated noise is flat, with a spectrum level equal to A_h , the average amount of energy passing through this binocular channel is

$$\frac{2\pi}{b^2} \left(A_h + A_g + \frac{bA_g}{c} e^{-\frac{\pi^2 X^2}{c^2}} e^{-\frac{\pi^2 Y^2}{c^2}} e^{-\frac{2b^2\epsilon_f^2\pi^2\sigma^2}{c^2}} \cos \left[\frac{2\pi\epsilon_f b^2 X}{c^2} + 2\pi\eta_f Y \right] \right). \quad (A1.7)$$

Thus, Equation A1.7 specifies the total amount of noise in the binocular channel centered at ϵ_f, η_f .

To compute the signal power in this binocular channel, we assume a sinusoidal grating of arbitrary rotation as the right-eye target. The equation describing such a grating is

$$A \cos[2\pi\epsilon_i x + 2\pi\eta_i y]. \quad (A1.8)$$

The left-eye grating is the identical grating vertically and horizontally shifted, where d_{tx} and d_{ty} specify the amount of horizontal and vertical shift in the target. Note that the extent of the horizontal and vertical shifts in the target need not be the same as the shifts applied to the noise background. The summated stimulus pattern in disparity channel $[d_\alpha, d_\beta]$ is

$$\begin{aligned} & A \cos[2\pi\epsilon_i x + 2\pi\eta_i y] + A \cos[2\pi\epsilon_i(x + d_{xt} + d_\alpha) + 2\pi\eta_i(y + d_{yt} + d_\beta)] \\ &= A \cos[2\pi(\epsilon_i x + \eta_i y)] + A \cos[2\pi(\epsilon_i x + \eta_i y) + 2\pi(\epsilon_i d_{tx} + \epsilon_i d_\alpha + \eta_i d_{ty} + \eta_i d_\beta)]. \end{aligned} \quad (A1.9)$$

Equation A1.9 shows that the luminance profile for the left-eye grating in the (d_α, d_β) disparity channel is simply a phase-shifted version of the luminance profile in the right-eye channel. Now, if we compute the average power along the axis of rotation of the sinusoidal grating, the average power in the summated target is

$$\frac{A^2}{2} [2 + 2\cos[2\pi(\epsilon_i d_{xt} + \epsilon_i d_\alpha + \eta_i d_{yt} + \eta_i d_\beta)]]. \quad (A1.10)$$

Again, note that, because of jitter, the amount of target displacement between the two eyes varies randomly, so that the power at any one moment is

$$\frac{A^2}{2} [2 + 2\cos[2\pi(\epsilon_i d_{xt} + \epsilon_i d_\alpha + \epsilon_i r + \eta_i d_{yt} + \eta_i d_\beta)]], \quad (A1.11)$$

where r is assumed to be a normally distributed random variable. The average value of the signal power, therefore, is

$$A^2 + A^2 e^{-2\pi^2\epsilon_i^2\sigma^2} \cos[2\pi(\epsilon_i d_{xt} + \epsilon_i d_\alpha + \eta_i d_{yt} + \eta_i d_\beta)]. \quad (A1.12)$$

Therefore, the ratio of signal power to noise power in disparity channel (d_α, d_β) centered at ϵ_i, η_i is

$$\frac{A^2 + A^2 e^{-2\pi^2 \epsilon_i^2 \sigma^2} \cos[2\pi(\epsilon_i d_{xt} + \epsilon_i d_\alpha + \eta_i d_{yt} + \eta_i d_\beta)]}{\frac{2\pi}{b^2} \left(A_h + A_g + \frac{bA_g}{c} e^{-\frac{\pi^2 X^2}{c^2}} e^{-\frac{\pi^2 Y^2}{c^2}} e^{-\frac{2b^2 \epsilon_i^2 \pi^2 \sigma^2}{c^2}} \cos \left[\frac{2\pi \epsilon_i b^2 X}{c^2} + 2\pi \eta_i Y \right] \right)} \quad (\text{A1.13})$$

For vertical gratings, $\eta_i = 0$, and, typically, $d_{yg} = k/(2\epsilon_i)$, $d_{yg} = 0$, $d_\beta = 0$, $d_{yt} = 0$ (k is an odd integer), for both binoptic and dichoptic conditions. In the dichoptic condition, $d_{xt} = 0$, and the maximum S/N ratio occurs in the ($d_\alpha = 0, d_\beta = 0$) disparity channel centered at $\epsilon_f = \epsilon_i$, $\eta_f = 0$. Therefore, the S/N ratio in the dichoptic condition is

$$\frac{A^2 + A^2 e^{-2\pi^2 \epsilon_i^2 \sigma^2}}{\frac{2\pi}{b^2} \left(A_h + A_g + \frac{bA_g}{c} e^{-\frac{\pi^2}{4c^2 k^2 \epsilon_i^2}} e^{-\frac{2b^2 \epsilon_i^2 \pi^2 \sigma^2}{c^2}} \cos \left[\frac{\pi b^2}{c^2} \right] \right)} \quad (\text{A1.14})$$

In the binoptic condition, $d_{xt} = k/(2\epsilon_i)$ for k , an odd integer, and the maximum S/N ratio occurs in the [$d_\alpha = -k/(2\epsilon_i), d_\beta = 0$] disparity channel. Therefore, the S/N ratio in the binoptic condition is

$$\frac{A^2 + A^2 e^{-2\pi^2 \epsilon_i^2 \sigma^2}}{\frac{2\pi}{b^2} \left(A_h + A_g + \frac{bA_g}{c} e^{-\frac{2b^2 \epsilon_i^2 \pi^2 \sigma^2}{c^2}} \right)} \quad (\text{A1.15})$$

The ratio of the S/N ratio in the dichoptic condition to that in the binoptic condition specifies the dichoptic advantage. This ratio is

$$\frac{A_h + A_g + \frac{bA_g}{c} e^{-\frac{2b^2 \epsilon_i^2 \pi^2 \sigma^2}{c^2}}}{A_h + A_g + \frac{bA_g}{c} e^{-\frac{\pi^2}{4c^2 \epsilon_i^2}} e^{-\frac{2b^2 \epsilon_i^2 \pi^2 \sigma^2}{c^2}} \cos \left[\frac{\pi b^2}{c^2} \right]} \quad (\text{A1.16})$$

Finally, when the external noise is much greater than the internal noise, so that we can assume $A_h = 0$, the dichoptic advantage becomes

$$\frac{1 + \frac{b}{c} e^{-\frac{2b^2 \epsilon_i^2 \pi^2 \sigma^2}{c^2}}}{1 + \frac{b}{c} e^{-\frac{\pi^2}{4c^2 \epsilon_i^2}} e^{-\frac{2b^2 \epsilon_i^2 \pi^2 \sigma^2}{c^2}} \cos \left[\frac{\pi b^2}{c^2} \right]} \quad (\text{A1.17})$$

Therefore, the predicted size of the BMLD becomes

$$\text{BMLD} = 10 \log \left[\frac{1 + \frac{b}{c} e^{-\frac{2b^2 \epsilon_i^2 \pi^2 \sigma^2}{c^2}}}{1 + \frac{b}{c} e^{-\frac{\pi^2}{4c^2 \epsilon_i^2}} e^{-\frac{2b^2 \epsilon_i^2 \pi^2 \sigma^2}{c^2}} \cos \left[\frac{\pi b^2}{c^2} \right]} \right] \quad (\text{A1.18})$$

Two features concerning the predicted size of the BMLD should be noted. First, if there is no eye jitter ($\sigma = 0$), then Equation A1.18 reduces to

$$10 \log \left[\frac{2}{1 - e^{-\frac{\pi^2}{4b^2 \epsilon_i^2}}} \right] \quad (\text{A1.19})$$

Now, if the bandwidth of the spatial-frequency channel (when expressed in octaves) is independent of frequency, it is easy to show that b is inversely proportional to the center frequency of the channel. Therefore, if there is no eye jitter and the bandwidth (when expressed in octaves) is independent of frequency, the size of the BMLD is independent of frequency.

Second, the primary effect of eye jitter is to decrease the size of the BMLD as frequency increases. In fact, it is easy to show that the size of the BMLD approaches zero as ϵ_i approaches infinity. Thus, the larger the size of σ , the more rapidly the BMLD approaches zero as frequency increases.

If we assume that the standard deviation of the eye jitter is about 2.1 min of arc and that the bandwidth of the spatial frequency channels is $\frac{1}{2}$ octave of their center frequency, the predicted size of the BMLD ranges from about 12 dB at 1 cycle/deg to about 6.5 dB at 4.5 cycles/deg. Moreover, the predicted values are quite close to the average values obtained in previous experiments at many different frequencies.

2. Template Matching Models of Phase Discrimination

Note that, in both the template-matching and the Bennett models, the unconstrained version of these models assumes that the observer always processes the whole target/foil area (72×72 pixels, or $4.12^\circ \times 4.12^\circ$). On the other hand, all the constrained versions of these models assume that the observer processes only the foveal portion of the target/foil area (32×32 pixels, or $2.06^\circ \times 2.06^\circ$).

Binoptic Conditions

Unconstrained observer. In a template-matching model, the observer convolves the template with the target/foil area of the visual display (72×72 pixels). Consider a single horizontal line along this display. The pixel luminance at location i along this line is given by

$$A \cdot \cos \left[2\pi \frac{(i + d_k)}{18} \right] + B \cdot \cos \left[2\pi n \frac{(i + d_k)}{18} + \theta \right] + z_i + 128, \quad (\text{A2.1})$$

where A was 13, d_k took on one of four values, n was 3 in Experiment 1 and 2 in Experiment 2, B was varied experimentally, the value of θ was $\pi/4$ for a target and $-\pi/4$ for a foil, and z_i was a normally distributed random variable with a mean of 0 and a standard deviation equal to 80. (The eight different templates are given by the four combinations of horizontal displacements and two values of θ . The equations for the templates are the same as Equation 2.1, without the normal random deviates.) The 72 random normal deviates describe a luminance pattern added to a template and, as such, can be decomposed in a sum of 36 sine and cosine components, so that the luminance equation for a line of the display, $L(i, n, d_k, \theta)$, when the target is added, becomes

$$\begin{aligned} L(i, n, d_k, \theta) &= \left[A \cdot \cos \left[2\pi \frac{(i + d_k)}{18} \right] + B \cdot \cos \left[2\pi n \frac{(i + d_k)}{18} + \theta \right] + \sum_{j=1}^{36} \left(a_j \cos \left[2\pi \frac{j}{72} i \right] + b_j \sin \left[2\pi \frac{j}{72} i \right] \right) + 128 \right] \\ &= \sum_{j=1}^{36} \left(\alpha_j \cos \left[2\pi \frac{j}{72} i \right] + \beta_j \sin \left[2\pi \frac{j}{72} i \right] \right) + 128, \end{aligned}$$

where $\alpha_j = a_j$, $\beta_j = b_j$; for $j \neq 4$ or $4n$,

$$\begin{aligned} \alpha_4 &= A \cdot \cos \left[2\pi \frac{d_k}{18} \right] + a_4, \\ \beta_4 &= -A \cdot \sin \left[2\pi \frac{d_k}{18} \right] + b_4, \\ \alpha_{4n} &= B \cdot \cos \left[2\pi n \frac{d_k}{18} + \theta \right] + a_{4n}, \\ \beta_{4n} &= -B \cdot \sin \left[2\pi n \frac{d_k}{18} + \theta \right] + b_{4n}, \end{aligned} \quad (\text{A2.2})$$

where the a and b terms are random normal deviates each having a mean of 0 and a standard deviation of 40/3. If we now convolve a line of the display with any one of the eight templates, $T(i, n, d_t, \theta_t)$, we obtain

$$\begin{aligned} &\int_{i=-36}^{36} T(i, n, d_t, \theta_t) \cdot L(i, n, d_k, \theta) di \\ &= \left[36\alpha_4 A \cdot \cos \left[2\pi \frac{d_t}{18} \right] - 36\beta_4 A \cdot \sin \left[2\pi \frac{d_t}{18} \right] + 36\alpha_{4n} B \cdot \cos \left[2\pi n \frac{d_t}{18} + \theta_t \right] - 36\beta_{4n} B \cdot \sin \left[2\pi n \frac{d_t}{18} + \theta_t \right] \right], \end{aligned} \quad (\text{A2.3})$$

where d_t refers to the horizontal shift employed in the template, and θ_t refers to the phase angle used in the template. Equation A2.3 shows that the cross-correlation of a template with a spatial display reduces to a linear combination of four independent random variables.

Note that Equation A2.3 is the cross-correlation for just 1 of the 72 lines of the spatial display. This means that there are 71 other cross-correlations for the same spatial display that differ only in that they have different sets of four independent random variables. An average across all 72 equations would be equivalent to computing just 1 correlation, but with the standard deviation of the normal deviates reduced by $72^{1/2}$. Therefore, Equation A2.3 gives the cross-correlation for the template and the pattern for n lines of the display, when the standard deviation of the normal variates is set to $(40/3)/n^{1/2}$. The efficiency of the noise reduction process (the percentage by which the noise is reduced over that contained in 1 line of the display is, therefore, $(n/72)^{1/2}$.

It should also be noted that an observer who uses n lines of length 72 will obtain the same degree of noise reduction as an observer who uses m lines of length m ($m < 72$), so long as $m^2 = 72n$. To see why this is the case, consider an observer who uses 36 lines of length 36 versus an observer who uses 18 lines of length 72. Now, take 18 of the 36 lines of length 36 and append them to the end of the remaining 36 lines. The result is 18 lines of length 72. In general, the degree of noise reduction is proportional to the square root of the number of pixels. Therefore, we can evaluate an observer whose effective field is 36×36 pixels by evaluating an observer who used 18 lines of 72 pixels.

On a simulated trial, one of the targets was chosen for the first interval and one of the foils for the second interval. Four normal random deviates (mean of 0, standard deviation = $(40/3)/n^{1/2}$) were generated for the first interval. Cross-correlating each of the templates with the target in the first interval produces eight equations of the form shown in Equation A2.3. The cross-correlated template producing the highest product was recorded for this interval. Four independent random deviates were generated for the second interval. The same eight templates were cross-correlated with the foil chosen in the second interval to produce eight cross-correlations. The template producing the highest value was again recorded. If the template producing the highest cross-correlation in the first interval was a target and the one producing the highest correlation in the second interval was a foil, a correct identification was recorded. If both intervals were labeled targets, a correct identification was scored if the cross-correlation in the first interval was greater than that in the second. If both intervals were labeled foils, a correct identification was scored if the cross-correlation in the second interval was greater than that in the first interval. A total of 30,000 trials were run to determine the probability of a correct identification for any particular condition.

Band-limited observer. The unconstrained observer convolves the template with each line of the display. The implication of this convolution is that the bandwidths of the spatial frequency filters tuned to ϵ and $n\epsilon$ are zero, so that the outputs of these filters at each frequency are precisely given by Equation A2.3. For a more realistic observer, the bandwidths of these filters would be considerably greater than zero and would differ for ϵ and $n\epsilon$. Therefore, the internal representation of each frequency would be inexact. To obtain coefficients of the sine and cosine terms at a particular frequency, we can convolve $L(i, n, d_k, \theta)$, with sine and cosine filters whose receptive fields are given by

$$\sin\left(2\pi \frac{m}{72} i\right) e^{-\gamma_m^2 i^2}, \cos\left(2\pi \frac{m}{72} i\right) e^{-\gamma_m^2 i^2}, \quad (\text{A2.4})$$

where γ_m is the parameter that determines the size of the Gaussian profiled receptive field and the bandwidth of the spatial-frequency filter for frequency $m/72$. Thus, the internal representation of a line of the display becomes

$$L_{\text{int}}(i, n, d_k, \theta) = \sum_{m=1}^{36} \left[\alpha'_m \cos\left(2\pi \frac{m}{72} i\right) + \beta'_m \sin\left(2\pi \frac{m}{72} i\right) \right],$$

where

$$\begin{aligned} \alpha'_m &= \int_{i=-36}^{36} L(i, n, d_k, \theta) \cos\left(2\pi \frac{m}{72} i\right) e^{-\gamma_m^2 i^2} di, \\ \beta'_m &= \int_{i=-36}^{36} L(i, n, d_k, \theta) \sin\left(2\pi \frac{m}{72} i\right) e^{-\gamma_m^2 i^2} di. \end{aligned} \quad (\text{A2.5})$$

If we now convolve one of the templates with this internal representation, all but the $m = 4$ and $m = 4n$ terms vanish, leaving

$$\begin{aligned} & \int_{i=-36}^{36} T(i, n, d_t, \theta_t) L_{\text{int}}(i, n, d_k, \theta) di \\ &= \left[36A \cos\left(2\pi \frac{d_t}{18}\right) \alpha'_4 - 36A \sin\left(2\pi \frac{d_t}{18}\right) \beta'_4 + 36B \cos\left(2\pi n \frac{d_t}{18} + \theta_t\right) \alpha'_{4n} - 36B \sin\left(2\pi n \frac{d_t}{18} + \theta_t\right) \beta'_{4n} \right], \end{aligned} \quad (\text{A2.6})$$

where

$$\begin{aligned}
 \alpha'_4 &= \frac{\sqrt{\pi}}{2\gamma_4} \left[\sum_{m=1}^{36} \alpha_m \left(e^{-\frac{\pi^2(m-4)^2}{\gamma_4^2 72^2}} + e^{-\frac{\pi^2(m+4)^2}{\gamma_4^2 72^2}} \right) \right] \\
 \beta'_4 &= \frac{\sqrt{\pi}}{2\gamma_4} \left[\sum_{m=1}^{36} \beta_m \left(e^{-\frac{\pi^2(m-4)^2}{\gamma_4^2 72^2}} - e^{-\frac{\pi^2(m+4)^2}{\gamma_4^2 72^2}} \right) \right] \\
 \alpha'_{4n} &= \frac{\sqrt{\pi}}{2\gamma_{4n}} \left[\sum_{m=1}^{36} \alpha_m \left(e^{-\frac{\pi^2(m-4n)^2}{\gamma_{4n}^2 72^2}} + e^{-\frac{\pi^2(m+4n)^2}{\gamma_{4n}^2 72^2}} \right) \right] \\
 \beta'_{4n} &= \frac{\sqrt{\pi}}{2\gamma_{4n}} \left[\sum_{m=1}^{36} \beta_m \left(e^{-\frac{\pi^2(m-4n)^2}{\gamma_{4n}^2 72^2}} - e^{-\frac{\pi^2(m+4n)^2}{\gamma_{4n}^2 72^2}} \right) \right].
 \end{aligned} \tag{A2.7}$$

Thus, using Equations A2.7 and A2.8, the results of a template match with a band-limited representation of a line of a display can be determined. Following the same procedures used for the unconstrained template-matching observer, the performance of a band-limited observer can be determined.

It should be noted that, since bandwidth is a function of frequency, the variance of the coefficients in Equation A2.7 will depend on frequency. If the bandwidth of the spatial frequency filters is constant in octave terms, the value of γ will increase with frequency, thereby producing a reduction in the variance of the coefficients as we go from $\epsilon + 2\epsilon$ to $\epsilon + 3\epsilon$. This implies that discrimination performance of the band-limited template matching should actually be better for $\epsilon + 3\epsilon$ than it is for $\epsilon + 2\epsilon$. This is exactly what we see in the simulations of Figure 10.

Dichoptic Conditions

Unconstrained observer. This dichoptic observer would sum a line from the left- and right-eye display and then convolve it with each of the templates. Recall that the right-eye noise display is displaced by 9 pixels. Therefore, the summed left- and right-eye displays become

$$\sum_{j=1}^{36} \left[\alpha_{d,j} \cos\left(2\pi j \frac{i}{72}\right) + \beta_{d,j} \sin\left(2\pi j \frac{i}{72}\right) \right],$$

where

$$\begin{aligned}
 \alpha_{d,j} &= a_j \left[1 + \cos\left(2\pi \frac{j}{8}\right) \right] + b_j \sin\left(2\pi \frac{j}{8}\right), j \neq 4 \text{ or } 4n \\
 \beta_{d,j} &= b_j \left[1 + \cos\left(2\pi \frac{j}{8}\right) \right] - a_j \sin\left(2\pi \frac{j}{8}\right), j \neq 4 \text{ or } 4n \\
 \alpha_{d,4} &= 2A \cos\left(2\pi \frac{d_k}{18}\right) + a_4 \left[1 + \cos\left(2\pi \frac{4}{8}\right) \right] + b_4 \sin\left(2\pi \frac{4}{8}\right) = 2A \cos\left(2\pi \frac{d_k}{18}\right) \\
 \beta_{d,4} &= -2A \sin\left(2\pi \frac{d_k}{18}\right) + b_4 \left[1 + \cos\left(2\pi \frac{4}{8}\right) \right] - a_4 \sin\left(2\pi \frac{4}{8}\right) = -2A \sin\left(2\pi \frac{d_k}{18}\right) \\
 \alpha_{d,4n} &= 2B \cos\left(2\pi n \frac{d_k}{18} + \theta\right) + a_{4n} \left[1 + \cos\left(2\pi \frac{4n}{8}\right) \right] + b_{4n} \sin\left(2\pi \frac{4n}{8}\right) \\
 \beta_{d,4n} &= -2B \sin\left(2\pi n \frac{d_k}{18} + \theta\right) + b_{4n} \left[1 + \cos\left(2\pi \frac{4n}{8}\right) \right] - a_{4n} \sin\left(2\pi \frac{4n}{8}\right).
 \end{aligned} \tag{A2.8}$$

Note that when $n = 3$, both the ϵ and 3ϵ components are completely unmasked. Therefore, when the summed pattern is convolved with one of the templates, the convolution will be unaffected by noise, and phase discrimination will be perfect. When $n = 2$, the convolution, however, will be affected by noise at 2ϵ . Therefore, performance will be less than perfect but better than in the binoptic case, where the convolution will be affected by noise at both ϵ and 2ϵ .

Band-limited and eye-jittered observer. Assume that the left- and right-eye images are added together, with some small degree of displacement, r . The equation for the sum of a single horizontal line presented to the left and the right eyes is

$$\begin{aligned}
& L_d(i, n, d_k, \theta) \\
&= \left[A \cdot \cos \left[2\pi \frac{(i+d_k)}{18} \right] + A \cdot \cos \left[2\pi \frac{(i+d_k+r)}{18} \right] + B \cdot \cos \left[2\pi n \frac{(i+d_k)}{18} + \theta \right] + B \cdot \cos \left[2\pi n \frac{(i+d_k+r)}{18} + \theta \right] \right. \\
&\quad \left. + \sum_{j=1}^{36} \left(a_j \cos \left[2\pi \frac{j}{72} i \right] + b_j \sin \left[2\pi \frac{j}{72} i \right] + a_j \cos \left[2\pi \frac{j}{72} (i+9+r) \right] + b_j \sin \left[2\pi \frac{j}{72} (i+9+r) \right] \right) + 128 \right] \\
&= \sum_{j=1}^{36} \left[\alpha_{d,j} \cos \left(2\pi \frac{j}{72} i \right) + \beta_{d,j} \sin \left(2\pi \frac{j}{72} i \right) \right], \tag{A2.9}
\end{aligned}$$

where

$$\begin{aligned}
\alpha_{d,j} &= \left[a_j \left(1 + \cos \left(2\pi \frac{j}{8} \right) \cos \left(2\pi j \frac{r}{72} \right) - \sin \left(2\pi \frac{j}{8} \right) \sin \left(2\pi j \frac{r}{72} \right) \right) \right. \\
&\quad \left. + b_j \left(\sin \left(2\pi \frac{j}{8} \right) \cos \left(2\pi j \frac{r}{72} \right) + \cos \left(2\pi \frac{j}{8} \right) \sin \left(2\pi j \frac{r}{72} \right) \right) \right], j \neq 4 \text{ or } 4n \\
\beta_{d,j} &= \left[b_j \left(1 + \cos \left(2\pi \frac{j}{8} \right) \cos \left(2\pi j \frac{r}{72} \right) - \sin \left(2\pi \frac{j}{8} \right) \sin \left(2\pi j \frac{r}{72} \right) \right) \right. \\
&\quad \left. - a_j \left(\sin \left(2\pi \frac{j}{8} \right) \cos \left(2\pi j \frac{r}{72} \right) + \cos \left(2\pi \frac{j}{8} \right) \sin \left(2\pi j \frac{r}{72} \right) \right) \right], j \neq 4 \text{ or } 4n \tag{A2.10}
\end{aligned}$$

and

$$\begin{aligned}
\alpha_{d,4} &= \left[A \left(1 + \cos \left(2\pi \frac{r}{18} \right) \right) \cos \left(2\pi \frac{d_k}{18} \right) - A \sin \left(2\pi \frac{d_k}{18} \right) \sin \left(2\pi \frac{r}{18} \right) \right. \\
&\quad + a_4 \left(1 + \cos \left(2\pi \frac{4}{8} \right) \cos \left(2\pi 4 \frac{r}{72} \right) - \sin \left(2\pi \frac{4}{8} \right) \sin \left(2\pi 4 \frac{r}{72} \right) \right) \\
&\quad \left. + b_4 \left(\sin \left(2\pi \frac{4}{8} \right) \cos \left(2\pi 4 \frac{r}{72} \right) + \cos \left(2\pi \frac{4}{8} \right) \sin \left(2\pi 4 \frac{r}{72} \right) \right) \right] \\
\beta_{d,4} &= \left[-A \left(1 + \cos \left(2\pi \frac{r}{18} \right) \right) \sin \left(2\pi \frac{d_k}{18} \right) - A \cos \left(2\pi \frac{d_k}{18} \right) \sin \left(2\pi \frac{r}{18} \right) \right. \\
&\quad + b_4 \left(1 + \cos \left(2\pi \frac{4}{8} \right) \cos \left(2\pi 4 \frac{r}{72} \right) - \sin \left(2\pi \frac{4}{8} \right) \sin \left(2\pi 4 \frac{r}{72} \right) \right) \\
&\quad \left. - a_4 \left(\sin \left(2\pi \frac{4}{8} \right) \cos \left(2\pi 4 \frac{r}{72} \right) + \cos \left(2\pi \frac{4}{8} \right) \sin \left(2\pi 4 \frac{r}{72} \right) \right) \right] \\
\alpha_{d,4n} &= \left[B \left(1 + \cos \left(2\pi n \frac{r}{18} \right) \right) \cos \left(2\pi n \frac{d_k}{18} + \theta \right) - B \sin \left(2\pi n \frac{d_k}{18} + \theta \right) \sin \left(2\pi n \frac{r}{18} \right) \right. \\
&\quad + a_{4n} \left(1 + \cos \left(2\pi \frac{4n}{8} \right) \cos \left(2\pi 4n \frac{r}{72} \right) - \sin \left(2\pi \frac{4n}{8} \right) \sin \left(2\pi 4n \frac{r}{72} \right) \right) \\
&\quad \left. + b_{4n} \left(\sin \left(2\pi \frac{4n}{8} \right) \cos \left(2\pi 4n \frac{r}{72} \right) + \cos \left(2\pi \frac{4n}{8} \right) \sin \left(2\pi 4n \frac{r}{72} \right) \right) \right] \\
\beta_{d,4n} &= \left[B \left(1 + \cos \left(2\pi n \frac{r}{18} \right) \right) \sin \left(2\pi n \frac{d_k}{18} + \theta \right) - B \cos \left(2\pi n \frac{d_k}{18} + \theta \right) \sin \left(2\pi n \frac{r}{18} \right) \right. \\
&\quad + b_{4n} \left(1 + \cos \left(2\pi \frac{4n}{8} \right) \cos \left(2\pi 4n \frac{r}{72} \right) - \sin \left(2\pi \frac{4n}{8} \right) \sin \left(2\pi 4n \frac{r}{72} \right) \right) \\
&\quad \left. - a_{4n} \left(\sin \left(2\pi \frac{4n}{8} \right) \cos \left(2\pi 4n \frac{r}{72} \right) + \cos \left(2\pi \frac{4n}{8} \right) \sin \left(2\pi 4n \frac{r}{72} \right) \right) \right]. \tag{A2.11}
\end{aligned}$$

Now, if each eye is filtered by spatial frequency channels of identical bandwidth, the effect is equivalent to passing the summated pattern through the same filter. Therefore, after passing through the spatial-frequency channels, the internal representation of the dichoptic stimulus is

$$L_{d,\text{Int}}(i, n, d_k, \theta) = \sum_{m=1}^{36} \left[\alpha'_{d,m} \cos\left(2\pi \frac{m}{72} i\right) + \beta'_{d,m} \sin\left(2\pi \frac{m}{72} i\right) \right]$$

where

$$\begin{aligned} \alpha'_{d,m} &= \int_{i=-36}^{36} L_d(i, n, d_k, \theta) \cos\left(2\pi \frac{m}{72} i\right) e^{-\gamma_m^2 i^2} di \\ \beta'_{d,m} &= \int_{i=-36}^{36} L_d(i, n, d_k, \theta) \sin\left(2\pi \frac{m}{72} i\right) e^{-\gamma_m^2 i^2} di. \end{aligned} \quad (\text{A2.12})$$

If we now convolve this internal representation of a binocularly summed stimulus with one of the templates, all frequencies except ϵ and $n\epsilon$ vanish, and the convolution becomes

$$\begin{aligned} \int_{i=-36}^{36} T(i, n, d_t, \theta_t) L_{d,\text{Int}}(i, n, d_k, \theta) di = \\ \left[36A \cos\left(2\pi \frac{d_t}{18}\right) \alpha'_{d,4} - 36A \sin\left(2\pi \frac{d_t}{18}\right) \beta'_{d,4} + 36B \cos\left(2\pi n \frac{d_t}{18} + \theta_t\right) \alpha'_{d,4n} - 36B \sin\left(2\pi n \frac{d_t}{18} + \theta_t\right) \beta'_{d,4n} \right] \end{aligned} \quad (\text{A2.13})$$

where

$$\begin{aligned} \alpha'_{d,4} &= \frac{\sqrt{\pi}}{2\gamma_4} \left[\sum_{m=1}^{36} \alpha_{d,m} \left(e^{-\frac{\pi^2(m-4)^2}{\gamma_4^2 72^2}} + e^{-\frac{\pi^2(m+4)^2}{\gamma_4^2 72^2}} \right) \right] \\ \beta'_{d,4} &= \frac{\sqrt{\pi}}{2\gamma_4} \left[\sum_{m=1}^{36} \beta_{d,m} \left(e^{-\frac{\pi^2(m-4)^2}{\gamma_4^2 72^2}} - e^{-\frac{\pi^2(m+4)^2}{\gamma_4^2 72^2}} \right) \right] \\ \alpha'_{d,4n} &= \frac{\sqrt{\pi}}{2\gamma_{4n}} \left[\sum_{m=1}^{36} \alpha_{d,m} \left(e^{-\frac{\pi^2(m-4n)^2}{\gamma_{4n}^2 72^2}} + e^{-\frac{\pi^2(m+4n)^2}{\gamma_{4n}^2 72^2}} \right) \right] \\ \beta'_{d,4n} &= \frac{\sqrt{\pi}}{2\gamma_{4n}} \left[\sum_{m=1}^{36} \beta_{d,m} \left(e^{-\frac{\pi^2(m-4n)^2}{\gamma_{4n}^2 72^2}} - e^{-\frac{\pi^2(m+4n)^2}{\gamma_{4n}^2 72^2}} \right) \right]. \end{aligned} \quad (\text{A2.14})$$

Thus, Equation A2.13 can be used in the simulations for a band-limited observer who sums the left and right patterns with momentary misalignment of the eyes (eye jitter). Of course, eye jitter has no effect on the binoptic condition, since, as target and background are displaced by the same amount, eye jitter would not affect the relative location of these two components of the stimulus.

3. Modified Versions of Bennett's Phase Discrimination Model

Binoptic Case

Unconstrained observer. The first step in Bennett's model is to determine the location of a peak in the lower spatial-frequency component of the compound grating. Before addition of the random noise, a single line of the pattern is given by

$$A \cdot \cos\left[2\pi \frac{(i+d_k)}{18}\right] + B \cdot \cos\left[2\pi n \frac{(i+d_k)}{18} + \theta\right] + 128. \quad (\text{A3.1})$$

A peak therefore occurs at location $i = -d_k$. When noise is added, however, the phase of this lower frequency component is changed. With the added noise, the equation for this frequency becomes (with α_4 and β_4 defined as in Equation A2.2)

$$\alpha_4 \cos\left[2\pi \frac{i}{18}\right] + \beta_4 \sin\left[2\pi \frac{i}{18}\right] = \sqrt{\alpha_4^2 + \beta_4^2} \cos\left[2\pi \frac{i}{18} + \omega\right], \quad (\text{A3.2})$$

where

$$\omega = \arctan \left[\frac{-\beta_4}{\alpha_4} \right].$$

Therefore, a peak at this spatial frequency is located at $i = -9\omega/\pi$.

A sine filter tuned to $n/18$ is then centered at $-9\omega/\pi$. This filter is convolved with a line of the pattern—that is, we determine

$$\int_{i=-36}^{36} L(i, h, d_k, \theta) \sin \left(2\pi \frac{4n}{72} \left[i + \frac{9\omega}{\pi} \right] \right) di. \quad (\text{A3.3})$$

In the absence of noise, the result of this convolution will be negative for any of the targets and positive for any of the foils. Therefore, the interval containing the smallest signed result of this convolution will be identified as the target.

In the Monte Carlo simulations, targets and foils were randomly chosen for the first and second intervals, and four independent random deviates were chosen for each of the two intervals. Again, the random deviates used in these simulations were assumed to have a mean of 0 and a standard deviation of $(40/3)/n^{1/2}$, where n is the number of lines of the display over which the filters are convolved. Thirty thousand trials were conducted for each datapoint. Noise reduction efficiency was defined as for the template model, above.

Band-limited observer. For the band-limited observer, the internal representation of ϵ (see Equation A2.10) is given by

$$\alpha'_4 \cos \left[2\pi \frac{i}{18} \right] + \beta'_4 \sin \left[2\pi \frac{i}{18} \right] = \sqrt{\alpha'^2_4 + \beta'^2_4} \cos \left[2\pi \frac{i}{18} + \omega' \right], \quad (\text{A3.4})$$

where

$$\omega' = \arctan \left[\frac{-\beta'_4}{\alpha'_4} \right].$$

A sine filter at 3ϵ centered at $-9\omega'/\pi$ is then convolved with the internal representation of the display to obtain a decision variable as for the unconstrained Bennett observer.

Dichoptic Case

Unconstrained observer. In the dichoptic case, the first step in the Bennett model is to locate a peak at frequency = $1/18$ for the sum of the left and right spatial displays. From Equation A2.8, we see that the sum of the left and right spatial displays for this frequency is

$$2A \cdot \cos \left[2\pi \frac{d_k}{18} \right] \cos \left[2\pi \frac{i}{18} \right] - 2A \cdot \sin \left[2\pi \frac{d_k}{18} \right] \sin \left[2\pi \frac{i}{18} \right]. \quad (\text{A3.5})$$

Therefore, following binocular summation, the phase of the ϵ component of the pattern is precisely determined so that a peak at this frequency can be located without error and a sine filter tuned to $n/18$ can be centered at $-d_k$. This filter is then convolved with the summed left- and right-eye displays. If the higher frequency is 3ϵ , Equation A2.8 shows that the error terms at 3ϵ will be eliminated so that this convolution will be carried out in the complete absence of noise. Thus, with summation of the left and right spatial patterns, the Bennett model demonstrates perfect performance. Summation, however, will not eliminate the noise at 2ϵ . Therefore, even in the absence of eye jitter, phase discrimination for ϵ and 2ϵ will not be perfect.

Band-limited and eye-jittered observer. For the band-limited and eye-jittered observer, the internal representation of the summed left- and right-eye patterns for frequency $1/18$ is

$$\alpha'_{d,4} \cos \left[2\pi \frac{i}{18} \right] + \beta'_{d,4} \sin \left[2\pi \frac{i}{18} \right] = \sqrt{\alpha'^2_{d,4} + \beta'^2_{d,4}} \cos \left[2\pi \frac{i}{18} + \omega'_d \right], \quad (\text{A3.6})$$

where

$$\omega'_d = \arctan \left[\frac{-\beta'_{d,4}}{\alpha'_{d,4}} \right].$$

Therefore, a sine filter (frequency = $n/18$) can be centered at $-9\omega'_d/\pi$ and convolved with the pattern. The result of this convolution is treated in the same way as was described in the unconstrained observer case of Bennett's model.