Vision Res. Vol. 32, No. 9, pp. 1695–1707, 1992
Printed in Great Britain.

0042-6989/92 $5.00 + 0.00

Lateral Interactions within Color Mechanisms in Simultaneous Induced Contrast

QASIM ZAIDI,* BILLIBON YOSHIMI,* NOREEN FLANIGAN,* ANTHONY CANOVA*

Received 5 November 1991; in revised form 11 February 1992

The perceived color of a region of visual space is a function not only of the spectral composition of the light incident from it, but also depends on the light incident from surrounding regions. The color contrast induced into a region is a result of lateral interactions between neural mechanisms. These interactions were studied by measuring the induced effect of circularly symmetric spatial sine-waves on a circular central test region. The phase of the surrounding sine-waves was changed uniformly in time, inducing a modulation in the appearance of the test. Observers adjusted the amplitude of real sinusoidal modulation in the test in order to null the induced modulation, and the nulling modulation was used as a measure of the induced effect. Spatial additivity was tested by using pairs of sine-waves of distinct spatial frequencies. The results showed that brightness induction can be characterized as a linear spatial process, i.e. the effects of parts of the surround at different distances from the test are summed, after the effect of each part is weighted by a negative exponential as a function of distance from the test. The magnitude of pure chromatic induction, however, is a result of nonlinear spatial interactions. Thus, these results have implications for the connections between visual mechanisms that process brightness and chromatic contrast.

Lateral interactions Spatial integration Spatial nonlinearity Induced contrast Color appearance

INTRODUCTION

The human visual system uses the spectral composition and radiance of incident light to compute perceived colors. This computation helps in segmenting visual scenes because different reflectances generally correlate with the phenomenal experience of different colors. Furthermore, the perceived color of a region of visual space is not just a function of the light incident from that region, but is also influenced by the light from adjoining regions in a manner that enhances the contrast between adjacent colors, and thus, aids the segmentation process (Chevreul, 1839). This phenomena is termed simultaneous induced color contrast. Besides segmentation, the colors of objects also help in other visual tasks, e.g. in locating and identifying a particular object in different situations. A special case is the relative invariance of the perceived colors in a scene despite changes in ambient illumination. A close connection exists between the lateral connections involved in induced contrast and some aspects of color constancy (e.g. Walvaren, 1976).

Regardless of any functional considerations, the phenomenal aspects of color induction provide an invaluable tool for studying the visual system. It is difficult, in general, to quantify absolute visual appearance. However, because the appearance of a test region can be altered by changing only the physical composition of the light incident from other regions in visual space, the change in appearance can be measured quantitatively by counteracting the induced change by a real change in the test region. The nature of the induced change can be used to infer properties of color processes. For instance Krauskopf, Zaidi and Mandler (1986) used this strategy, in conjunction with modulation along theoretically defined color directions, to show that the computation of color appearance involves higher level color mechanisms beyond the traditional opponent stage. As another instance, the present study demonstrates that quantifying the induced effect of appropriate spatially complex surrounds can reveal the nature of the lateral interactions involved in color computations.

This study explored the organization of lateral interactions within the three cardinal mechanisms of color vision (Krauskopf, Williams & Heeley, 1982). To this end, during each measurement, the colors of the surrounding regions were restricted to one of the cardinal directions of color space and systematically arranged in a circularly symmetric sinusoidal pattern. The magnitude of induced contrast was measured by a modulation nulling method. The results showed that for achromatic brightness induction, the combined effects of the surrounding regions could be characterized by a simple weighted integration. However, the results for pure chromatic induction indicated the presence of nonlinear lateral interactions.

*Department of Psychology, Columbia University, New York, NY 10027, U.S.A.

EXPERIMENT 1. SIMPLE SINUSOIDAL SURROUNDS

The differential effects of parts of the surround at different distances from the test were studied by using stimuli similar to those shown in Fig. 1. In Fig. 1, the central disks are the test regions. The surrounds look like blurred bulls-eye targets. Along every radial line through the center, the color of the surround varies as a spatial sine-wave along a color line through mid-white. Each of the top, middle, and bottom rows of Fig. 1 show surrounds consisting of circularly symmetric spatial sine-waves of a single frequency windowed by the edges of the surround. These patterns are termed circularly symmetric to denote that their shape is invariant to all rotations through the center. Each row shows three different phases of the sine-wave with respect to the inner edge of the surround. The central disks in all nine pictures are physically the same mid-white, but appear to have different brightnesses. Two aspects of the phenomenal appearance of the central test are directly relevant to the present study. First, within each row, as the phase of the surrounding sine-wave changes, the appearance of the test changes. The appearance of the test is roughly complementary to the appearance of the inner edge of the surround. Second, the magnitude of the change in the appearance is least in the top row which has the surround with the highest spatial frequency, and largest in the bottom row which has the surround with the

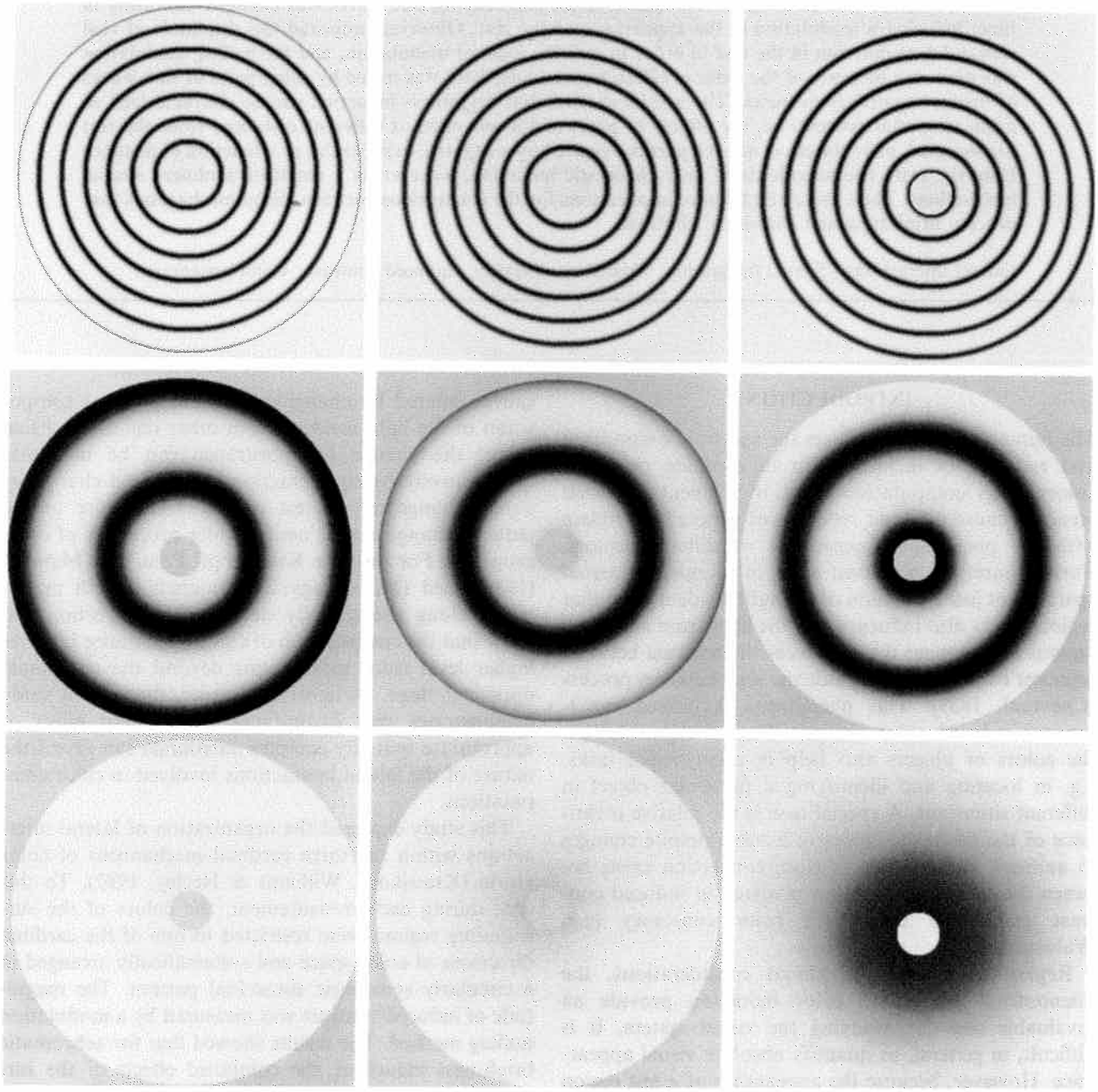


FIGURE 1. Top row shows a circular mid-white test disk surrounded by a circularly symmetric high frequency sine-wave in three different phases. Middle and bottom rows show the same test surrounded by sine-waves of medium and low spatial frequencies respectively. All test disks are physically similar; different apparent brightnesses are due to different amounts of induced contrast from the surrounds. No attempt was made to compensate for any nonlinearities in the photographic process.

lowest spatial frequency. Each row in Fig. 1 illustrates the procedure used in the first experiment. The surround consisted of a single sine-wave of one of seven different spatial frequencies and its color varied along one of the three cardinal directions of color space. As the phase of the surround (with respect to the inner edge) was changed uniformly in time, so that the sine-wave appeared to drift toward the center at a constant velocity, the appearance of the center changed cyclically in time. The induced modulation was nulled by the addition of real modulation to the test field, and the nulling modulation was used as the measure of the induced effect.

Stimulus parameters

The test was a spatially uniform disk, centered on the fovea, with a diameter subtending a visual angle of 1 deg. The inner edge of the surround coincided with the outer edge of the test. The diameter of the circular outer edge of the surround subtended 8 deg of visual angle. The magnitude of induction was measured for sine-waves with spatial frequencies (along radial lines) of 0.05, 0.1, 0.2, 0.4, 0.8, 1.6, and 3.2 c/deg. Each sine-wave was drifted inside the window towards the center at a speed of 1 Hz.

The color of the spatial sine-wave varied along one of the three cardinal directions of color space (Krauskopf *et al.*, 1982). The three dimensional color space is shown in Fig. 2 (Sachtler & Zaidi, 1991). The origin and end-points of the axes are described in approximate color names and in L-, M- and S-cone excitations (Smith & Pokorny, 1975) using the chromaticity coordinates of the MacLeod and Boynton (1978) diagram. The center of the space was metameric to white at 50 cd/m^2. Along the "light–dark" axis, the excitation of the three cone types was modulated proportionally with the same sign. Along the "green–red" axis, the excitation of L- and M-cones changed in equal and opposite amounts to keep their sum constant, and the excitation of S-cones was constant. Along the "yellow–violet" axis, only the excitation of S-cones changed. Along each axis, the color of the surround was modulated symmetrically, around the mid-white center of the space, to the ends of the axes shown in Fig. 2.

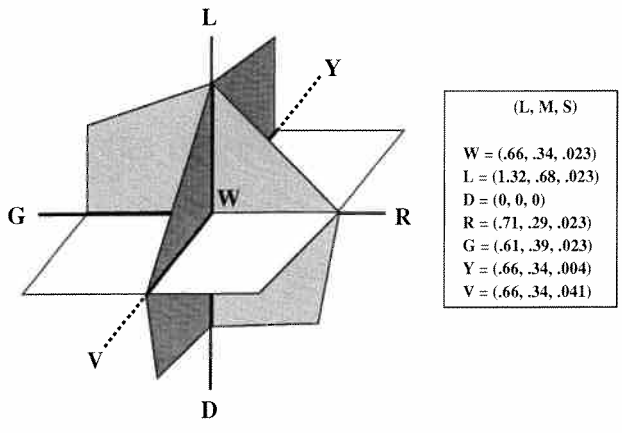


FIGURE 2. The three-dimensional space used for color specification, taken from Sachtler and Zaidi (1991).

The circular surround was enclosed within a spatially uniform, steady, 10.67 × 10 deg rectangle. The color of the outer rectangle was metameric to the center of the color space.

Measurement procedure

The modulation nulling technique discussed by Krauskopf *et al.* (1986) and Zaidi, Yoshimi and Flannigan (1991), was used to measure the amount of induction within the central test. When the sine-wave in the surround was drifted towards the center (1 Hz), a perceived modulation was induced into the test. The observer adjusted the amplitude of real sinusoidal modulation in the central disk to minimize the perceived modulation. The nulling modulation was along the same color line as the inducing modulation, had the same temporal frequency of 1 Hz and was in the same temporal phase as the inner edge of the surround. The amplitude of nulling modulation as a fraction of inducing amplitude was used as the estimate of the magnitude of induced contrast.

Each session consisted of 30 nulling measurements for a particular spatial configuration of the surround, 10 for each cardinal direction. The color directions of the trials were randomly interleaved within a session, and different spatial configurations were run in random order.

The nulling method was devised to overcome the limitations of matching and memory based measurements of simultaneous induction. [For a detailed explanation of these limitations see pp. 264–271 of Helmholz (1924).] The nulling method has the following advantages. (1) During each trial, the observer fixates on the central disk. Successive contrast effects are kept to a minimum, because, except for small eye movements, the portion of the retina that receives the induced modulation is never exposed to the inducing field. (2) The observer is required only to signal the absence or presence of modulation without extracting any subjective qualities of the test. At the null point, the induced effect of all phases of the surround modulation are equally canceled. No memory demands are made. The nulling method thus meets the requirements for Brindley's (1960) class A experiments. (3) In every trial, the chromaticity and luminance of every point of the display, averaged over a cycle of modulation, was metameric to the center of the color space in Fig. 2. The steady-state adaptation was therefore constant throughout the experiments.

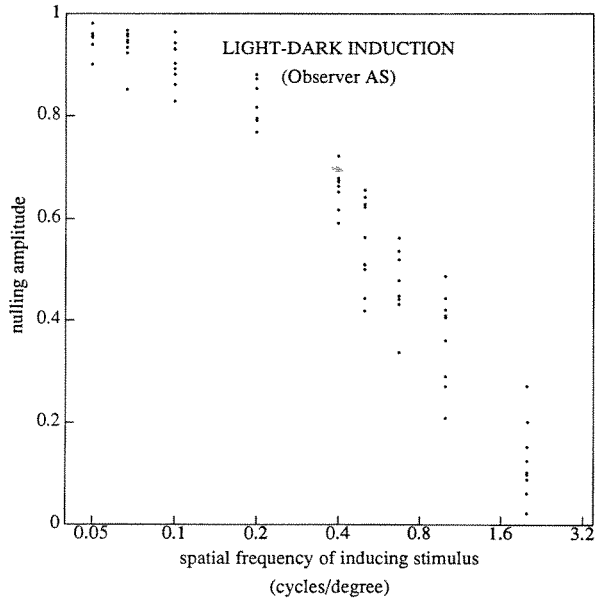
Krauskopf *et al.* (1982) and Shapiro and Zaidi (1991) have shown that prolonged viewing of temporally modulated fields has a significant effect on the response of color mechanisms. In the present experiments, these habituation effects were minimized by restricting the duration of each trial and by randomly interleaving trials along different color lines within a session. In addition, analysis of the data confirmed that measurements made in the first and second half of each session were not statistically different.

Equipment and stimulus generation

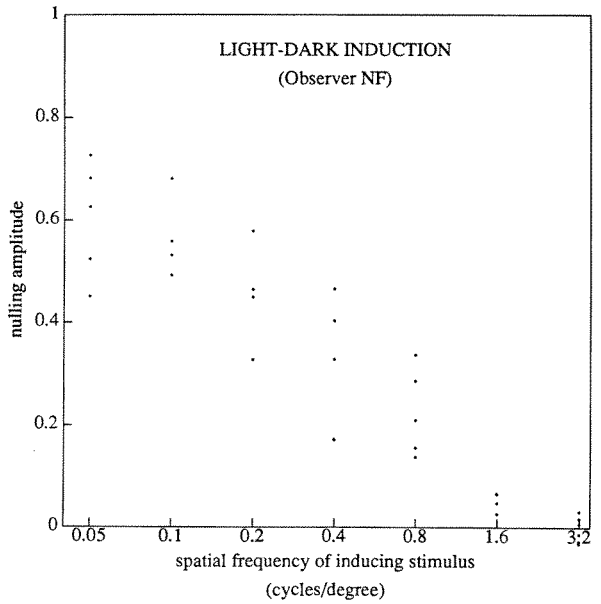
Stimuli were displayed on the screen of a Tektronix 690SR color monitor refreshed at the rate of 120 interlaced frames/sec. Images were generated using an Adage 3000 raster-based frame-buffer generator. 1024 output levels were specifiable for each gun. Up to 256 colors could be displayed on the screen in any frame. The color of an image was modulated sinusoidally in time by cycling through a pre-computed color array. Up to 240 colors could be changed during the fly-back interval in a screen refresh. The 512 × 480 pixel display subtended an angle of 10.67 × 10 deg at the observer's eye. Details of the calibration and color specification can be found in Zaidi and Halevy (1992). All stimulus presentation and data collection was done under automatic computer control.

Results

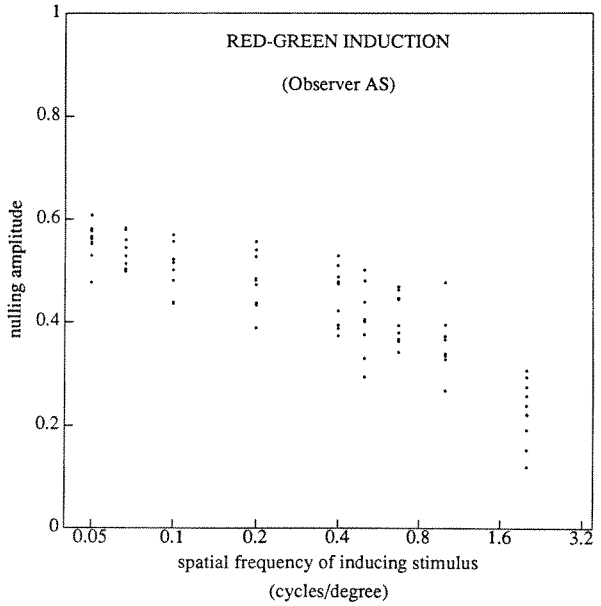
The relationship of principal interest in this experiment was between the magnitude of induced contrast in the test and the spatial pattern of the surround. The amplitude of nulling modulation as a fraction of the amplitude of inducing modulation is plotted versus the spatial frequency of the surrounding sine-wave (on a logarithmic scale) in Fig. 3. Results are "light–dark", "red–green" and "yellow–violet" induction are shown in Fig. 3 (a, b, c) respectively for two color normal observers. Each data point is the mean of 10 repeated observations. The results are very simple: for all three color directions, the magnitude of induction is greatest for the lowest spatial frequency and decreases monotonically with increasing spatial frequency of the surround.



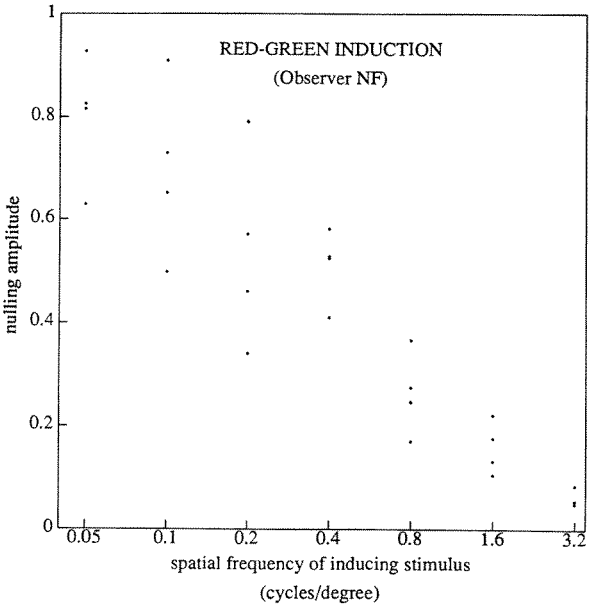
(a)



(a)



(b)



(b)

FIGURE 3. (a,b) *Caption on facing page.*

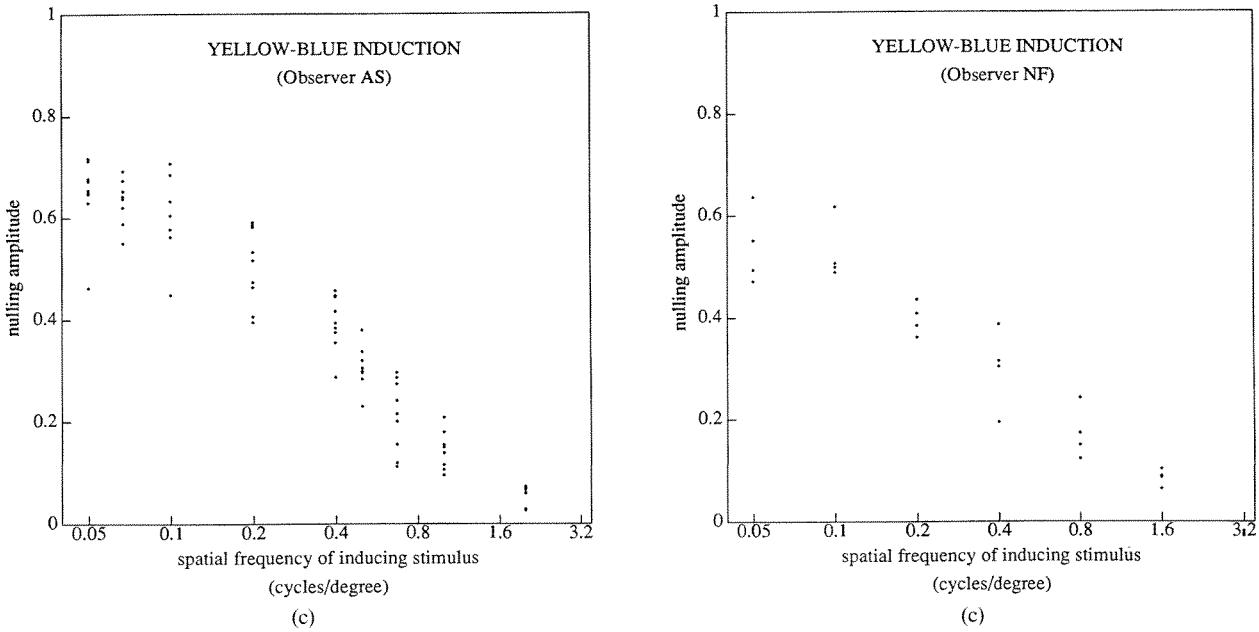


FIGURE 3. Amplitude of nulling modulation (as a fraction of inducing amplitude) plotted vs the spatial frequency of the surrounding sine-wave (on a logarithmic scale) for (a) "light–dark"; (b) "red–green"; and (c) "yellow–violet" induction, for two observers.

It will be shown in a later section that these low-pass "nulling" functions would be of considerable theoretical utility if a simple weighted-integration model were sufficient to explain the magnitude of induced contrast. Briefly, if it could be assumed that the amplitude of induced contrast was a weighted sum of the effect of each point of the surround and that the effect of each point was proportional to the contrast amplitude at that point, then the spatial weighting function for a surround of given size could be estimated from the Fourier transform of the measured "nulling" function. In more abstract terms, if the magnitude of induced contrast were a continuous linear functional of the spatial contrast of the surround, then all the information required to characterize the influence of the surround would be contained in the response to the appropriate sinusoidal stimuli. The issue of spatial additivity was the subject of the second experiment.

EXPERIMENT 2. COMPOUND SINUSOIDAL SURROUNDS

The additivity of the effect of different parts of the surround on the appearance of a test was studied by testing whether the induced effect of surrounds consisting of pairs of circularly symmetric sine-waves was equal to the sum of the induced effects of the constituent sine-waves. In Fig. 4, the central disks are the test regions. Along every line through the center, the surround consists of the sum of two spatial sine-waves whose colors vary along a color line around mid-white. Each of the top, middle and bottom rows show surrounds consisting of the sum of two spatial frequencies windowed by the edges of the surround. The three rows consist of the same medium frequency paired with a high (top), medium (middle) and low frequency (bottom). In all nine pictures, the two sine-waves are set to have the same phase at the inner edge of the surround. The central disks are physically the same mid-white, but appear to have different brightnesses depending on the phases of the sine-waves. The magnitude of the induced change in Fig. 4 also appears to depend on the frequency of the sine-wave paired with the medium frequency, being least with the paired high frequency and most with the paired low frequency.

Each row in Fig. 4 illustrates the procedure used in the second experiment. Sine-waves of each of the seven spatial frequencies used in Expt 1, were paired with another sine-wave of one of the seven frequencies, to give a total of 49 compound surrounds. The color of both constituents of each compound surround varied along the same cardinal direction. The amplitude of each constituent sine-wave was half that in Expt 1. Both constituent sine-waves were drifted towards the center at 1 Hz, so that they were always in the same phase at the inner edge of the surround. The inner edge of the surround therefore had a sinusoidal temporal modulation of 1 Hz with an amplitude equal to the surround modulation in Expt 1. The nulling modulation was along the same color line as the inducing modulation, had the same temporal frequency of 1 Hz, and was in the same temporal phase as the inner edge of the surround.

Results

The main question to be answered by this experiment was: does adding a sine-wave of a particular frequency to another sine-wave have a constant effect on the magnitude of induced contrast irrespective of the frequency of the paired sine-wave? The data from this experiment have a three-dimensional structure—there is a mean nulling modulation for each of the 7×7 spatial frequency pairs. However, it is easier to visualize

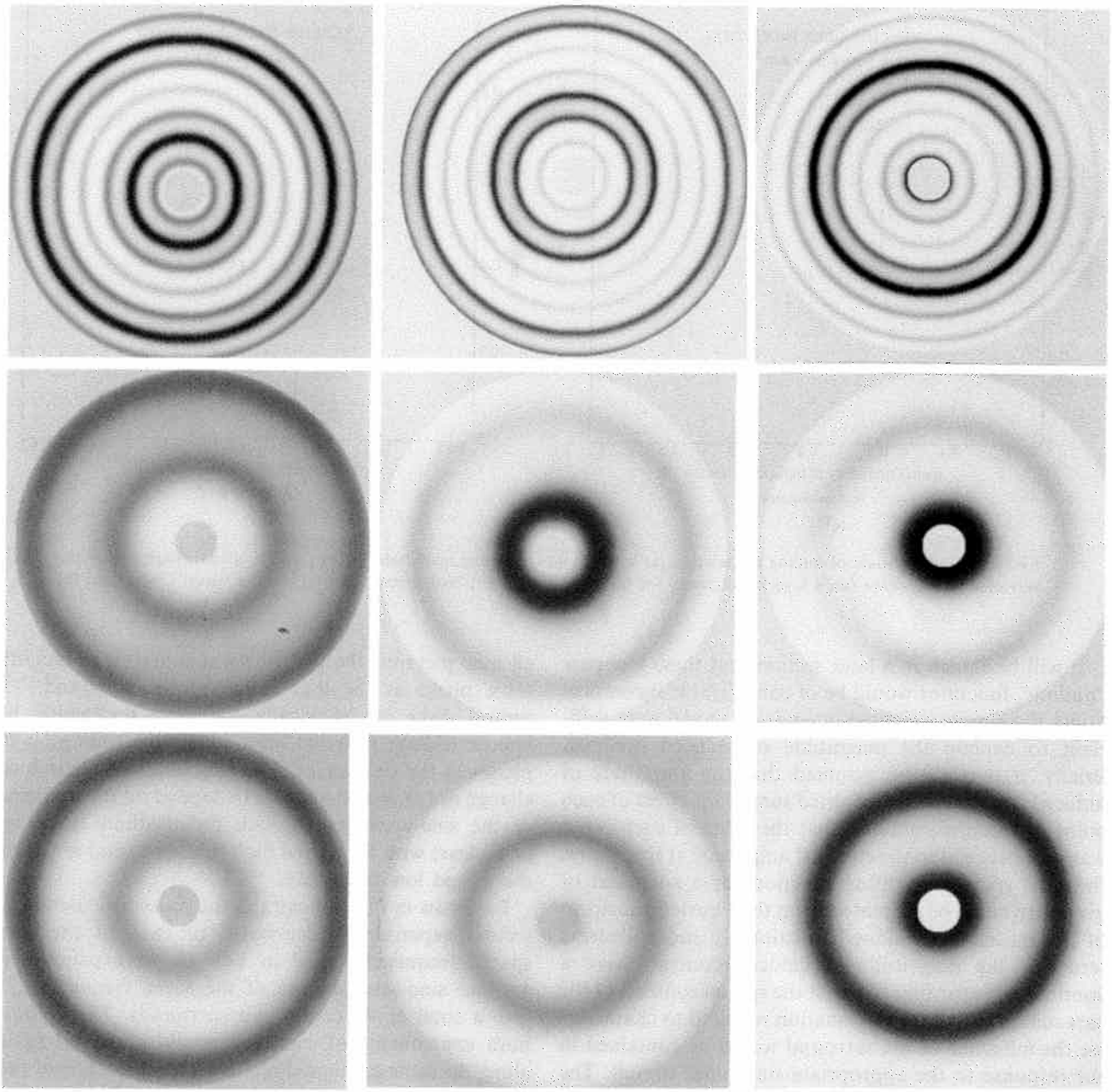


FIGURE 4. Circular mid-white test disks surrounded by circularly symmetric compound sine-waves in different phases. Top row shows a high frequency sine-wave added to a medium frequency sine-wave; middle row, two different middle frequency sine-waves added; and bottom row, a low frequency sine-wave added to a medium frequency. All test disks are physically the same, any difference in appearance is due to induced contrast from the surround.

additivity vs non-additivity in a 2-D diagram like Fig. 5 which has several cross-sections of the imaginary 3-D graph collapsed into one graph. Each of the graphs in Fig. 5 shows the mean data for the 49 compound stimuli for a particular cardinal color direction. Each point is the mean of 30 observations. The ordinate of each point is the amplitude of nulling modulation as a fraction of the amplitude of inducing modulation. Each piece-wise linear curve in the figure connects the data for a particular spatial frequency when paired with the spatial frequencies corresponding to the abscissa. Different line types have been used to distinguish the curves. The key to identifying the curves is to start at the left-most point where the curves are ordered from top to bottom in the same order as the spatial frequencies from 0.05 to 3.2 c/deg shown in the inset.

All the piece-wise linear curves for "light–dark" induction in Fig. 5(a) have roughly the same shape: the amount of modulation required to null the induced effect decreases as each frequency is paired with progressively higher frequencies. In addition, the curves are roughly parallel to each other, indicating that the incremental induced effect of adding sine-waves of different spatial frequencies (represented by the curves) is fairly independent of the paired spatial frequency (represented by the abscissa). The main departure from parallelism is the noticeable upward protuberance in each curve where the two paired frequencies were equal, i.e. the surround

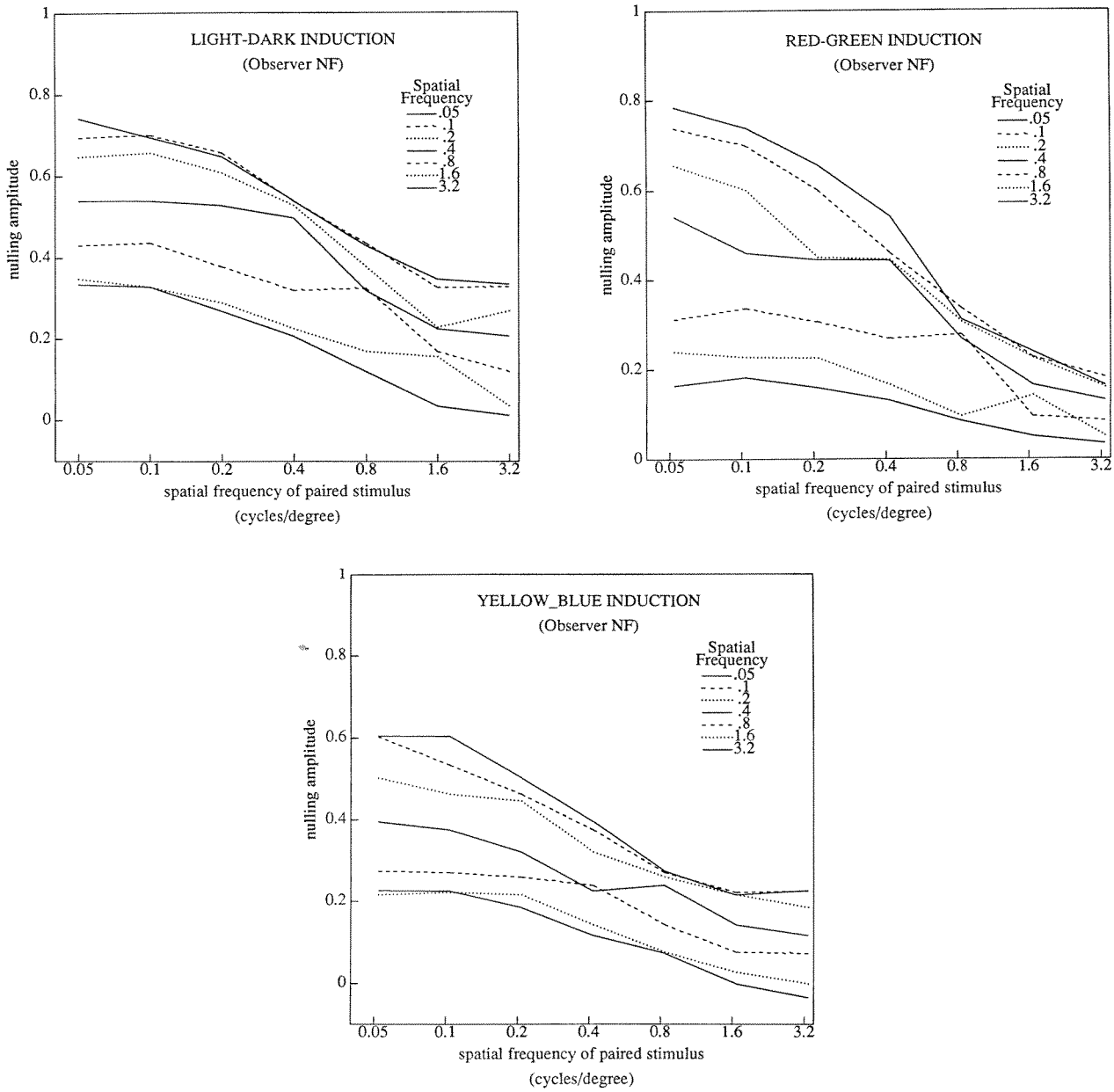


FIGURE 5. Amplitude of nulling modulation for each of the compound inducing stimuli used in Expt 2. The first component frequency of the surround is indicated by the line-type of the curve, and the second by the abscissa.

consisted of a single sine-wave. At other pairings, the amplitude required to null the effect of a compound surround is a little less than the sum of the amplitudes required to null the effects of the two components. As discussed later, this reduction could be due to phase-shifts in the induced modulation relative to the surround modulation, which vary with spatial frequency of the surround. The slight sub-additivity in nulling amplitude of the compound surround could therefore be due to a phase mismatch in the modulation induced by the components.

The results for "red–green" and "yellow–violet" induction in Figs 5(b, c), however, show significant departures from parallelism. The curves are spread apart towards the left (lower paired frequencies), and bunched together towards the right (higher paired frequencies). The effect of adding a high spatial frequency chromatic component severely attenuates the magnitude of induced contrast, especially for the "red–green" condition.

WEIGHTED SPATIAL INTEGRATION OF INDUCED "LIGHT–DARK" CONTRAST

The results of the "light–dark" conditions in Expt 1 and 2 can be interpreted in terms of a simple model that postulates a weighted spatial integration of induced contrast. The perceived color at a point in visual space has two components, one due to the radiance and spectral composition of the light emanating from that point and the second due to the total induced effect of surrounding points. The model makes three assumptions about the induced effect. First, if the colors in a display are restricted to one line in color space, the induced effect of any surrounding point is in the complementary

direction from the surround color with respect to the test color, with a magnitude proportional to the difference between the two colors. Second, the induced effect of each surrounding point is weighted by a decreasing function of spatial distance from the test point. Third, the total induced effect is simply the sum of the induced effects of individual surrounding points. Algebraically, the model is defined by equation (1):

$$y = -\int_{\Omega=0}^{2\pi} \frac{\int_0^{\infty} g(s)A(\Omega, s)\mathrm{d}s}{2\pi} \mathrm{d}\Omega \qquad (1)$$

where y is the total induced effect at the test point; Ω is the direction away from the test point; s is the spatial distance in units of deg of visual angle from the test, so that the test point is at $s = 0$; $g(s)$ is the monotonically decreasing spatial weighting function of s; and $A(\Omega, s)$ is the signed magnitude of the color difference between (Ω, s) and the test point.

The stimuli used in these experiments were circularly symmetric and varied only along radial lines. Therefore, by assuming isotropy of induced effects, equation (1) can be reduced to one direction radiating out from the center of the test field. For a drifted single sinusoid of spatial frequency equal to ϕ_i c/deg, the induced effect at time t for the center point of the circular test can be expressed as:

$$y(t, \phi_i) = -A\int_L^X g(s)\cos[2\pi(\rho_0 t - \phi_i s + \phi_i L)]\mathrm{d}s \qquad (2)$$

where A is the amplitude of the surround sine-wave; L is the inner edge of the surround, i.e. the radius of the test disk; X is the outer edge of the surround; and ρ_0 the temporal frequency of the drift (c/sec). Since the light emanating from every point in the test region is physically identical, in terms of the present model, the induced effect of every other point inside the test on the center is zero. Therefore equation (2) is expressed solely in terms of the effect of the surround.

Since the quantity measured in Expt 1 was the amplitude of the temporal modulation of the induced effect and not the instantaneous induction at time t, both sides of equation (2) were Fourier transformed into the temporal frequency domain. By exploiting the fact that the drift was at a constant velocity given by ρ_0 divided by ϕ_i, the Fourier transform of the right-hand-side was simplified to the expression shown in equation (3):

$$Y(\rho, \phi_i) = -\frac{A}{2}[\delta(\rho - \rho_0) + \delta(\rho + \rho_0)]e^{i2\pi\phi_i L}\int_L^X g(s)e^{-i2\pi\phi_i s}\mathrm{d}s \qquad (3)$$

where $Y(\rho, \phi_i)$ is the induced modulation for a surround of spatial frequency ϕ_i drifted toward the test point at a temporal frequency of ρ_0, and δ is the Dirac delta function. A detailed description of the derivation is presented in the Appendix. For each spatial frequency ϕ_i, the right-hand-side of equation (3) is the product of two expressions. The expression outside the integral consists of the temporal Fourier transform of a cosine with a temporal frequency equal to ρ_0, multiplied by a constant whose value depends on the radius of the test disk. The integral is equal to the spatial Fourier transform of $g(s)$ evaluated in the interval $[L, X]$, and has no temporal component. Equation (3) shows that, as a consequence of the linearity of the model, the temporal component of the induced contrast is a sinusoidal modulation with a frequency equal to ρ_0.

If the nulling modulation $N(\rho_0, \phi_i)$ for a single sinusoid is taken as an estimate of $-Y(\rho_0, \phi_i)$, then the curves in Fig. 3 describe the amplitude of induced temporal modulation as a function of the spatial frequency of the surround. If the three assumptions of the model are satisfied for a particular color line, then given the proper choice of $g(s)$, equation (3) will fit the data. Conversely, for a number of spatial frequencies, the right-hand-side can be equated to the empirical estimate of N, and $g(s)$ the weighting function can be estimated from the inverse spatial Fourier transform of N as a function of ϕ. However, the inverse transform requires both the amplitude and phase of N, whereas in these experiments only the amplitude was measured. Moreover, the induced amplitude was measured for only a small number of spatial frequencies, which would limit the resolution with which $g(s)$ could be estimated from the inverse transform. Therefore, a different strategy was adopted. Since many smooth monotonic functions can be approximated by exponential functions, it was assumed that the spatial weighting function could be approximated by a negative exponential function:

$$g(s) = \kappa e^{-\alpha s}. \qquad (4)$$

Equation (4) was substituted into equation (3), and expressions were derived for the amplitude and the temporal phase of induced modulation as a function of spatial frequency. These expressions and their derivation are shown in the Appendix [equations (A9) and (A10)].

For "light–dark" induction, the approximately parallel curves in Fig. 5(a) show that the spatial additivity assumptions of the model are roughly satisfied. In addition, Krauskopf *et al.* (1986) showed that for "light–dark" induction, the nulling amplitude is a linear function of the inducing amplitude. Therefore, the derived expression for the amplitude of N was fitted to the mean of the single sinusoid data using the Nelder–Mead simplex algorithm, and values of κ and α were estimated for the two observers. The best fitting curves are shown with the mean data in Fig. 6. The generally good fit of the curve to the data justifies the choice of a negative exponential for the weighting function. The spatial weighting factor based on the best estimates of κ and α is shown in Fig. 7 as a function of s, the distance from the center point. The relative value of the scaling parameter κ reflects the relative magnitude of the induced effect across observers and color directions. The parameter α is the space constant which turned out to be virtually identical for the two observers.

Examining the form of the spatial weighting functions in Fig. 1 gives an intuitive feel for the generation of the

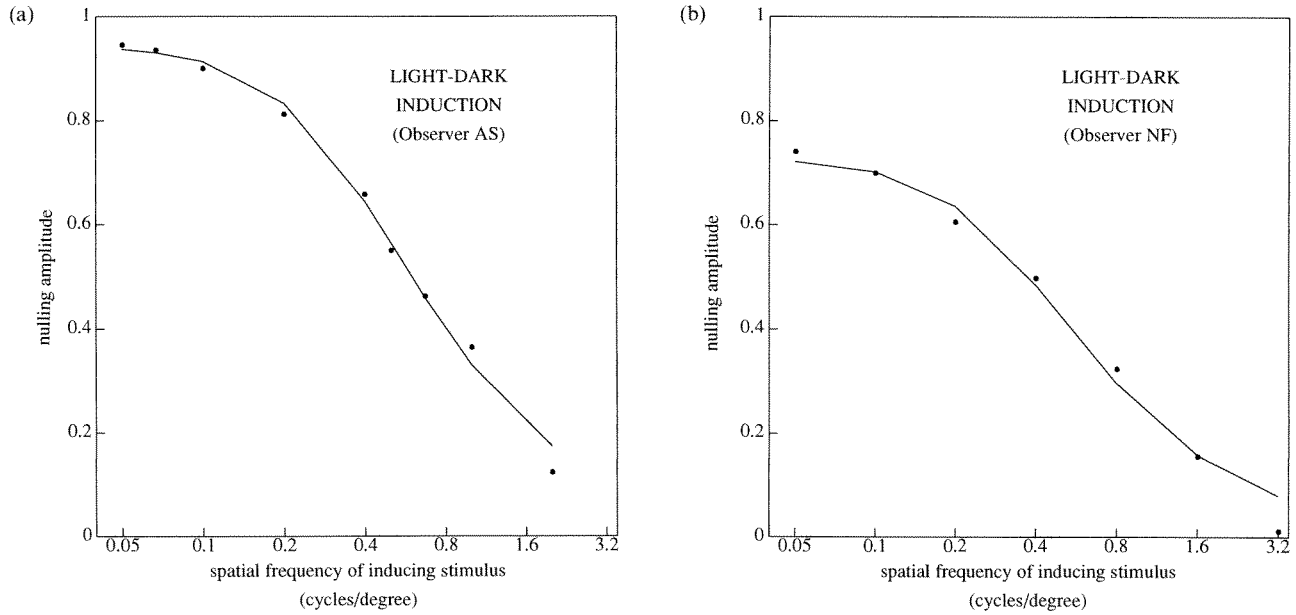


FIGURE 6. Amplitudes of induced temporal modulation, predicted from the negative exponentially weighted spatial integration model (solid line), fitted to the mean data for "light–dark" induction (symbols).

low-pass "light–dark" curves in Fig. 3. When the phase of a low-frequency surround is such that the inner edge is at the "lightest" point, the "light" part of the grating is multiplied by much larger numbers than the "dark" part, whereas when a high-frequency surround is in that phase, the "light" and "dark" parts are multiplied by numbers that are closer in magnitude. Hence a larger nulling amplitude is predicted for a low-frequency surround than a high-frequency surround. The steep slope of the weighting function indicates that local induced contrast at the edge of the test is a dominant factor in "light–dark" induction.

In a similar fashion, it can be visualized, that for surrounds that are not spatially uniform, the modulation induced into a steady mid-white test will not be in exact counterphase to the temporal modulation of the inner edge of the surround. If a spatially uniform surround is modulated purely temporally, then when the color of the inner edge of the surround (which is identical to the color of the whole surround) is at the mean level, i.e. the surround is the same color as the steady mid-white test, the weighted effect of the surround will be equal to zero, and the appearance of the test will not change from the mid-white. In addition, when the color of the surround is at one of the peak phases, the apparent color of the test will be at the opposite peak of the induced modulation. Therefore, for a uniform surround, the model predicts that the induced modulation will be in counterphase to the surround modulation. However, if the surround consists of a spatial sine-wave, when the inner edge is at the mean level, because $g(s)$ also gives weight to other parts of the surround, the weighted effect of the surround will not be zero. Similarly, the maxima of the induced modulation will not coincide in time with the peak phases of the modulation of the inner edge of the surround. Thus, the predicted induced modulation will not be in exact counter-phase to the modulation of the inner edge. Fig. 8 shows the predicted phase lags for one observer as a function of the spatial frequency of the inducing stimulus. A phase-lag equal to π radians denotes an induced modulation in exact counter-phase to the temporal modulation of the inner edge of the surround.

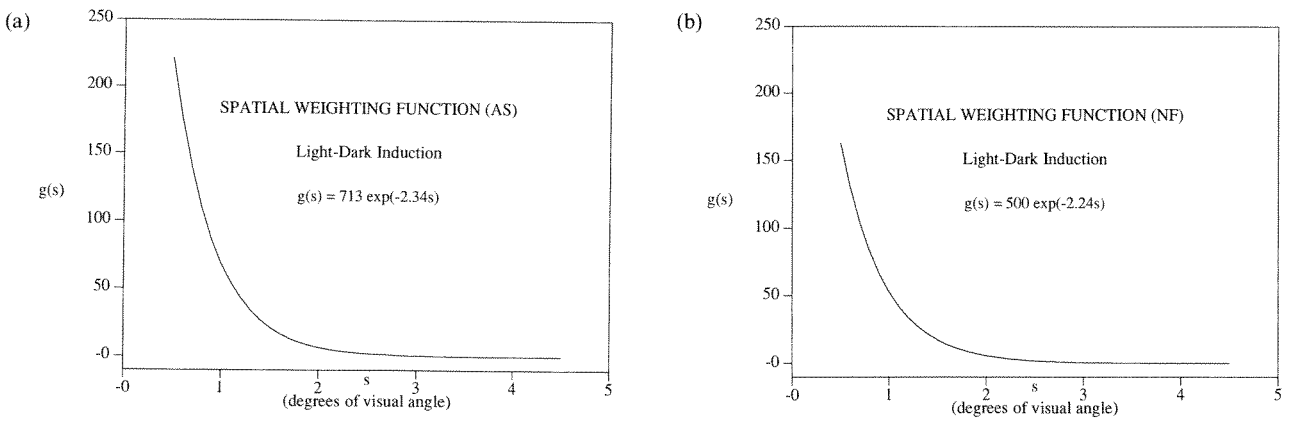


FIGURE 7. The spatial weighting function $g(s)$ for two observers plotted vs distance from the test point.

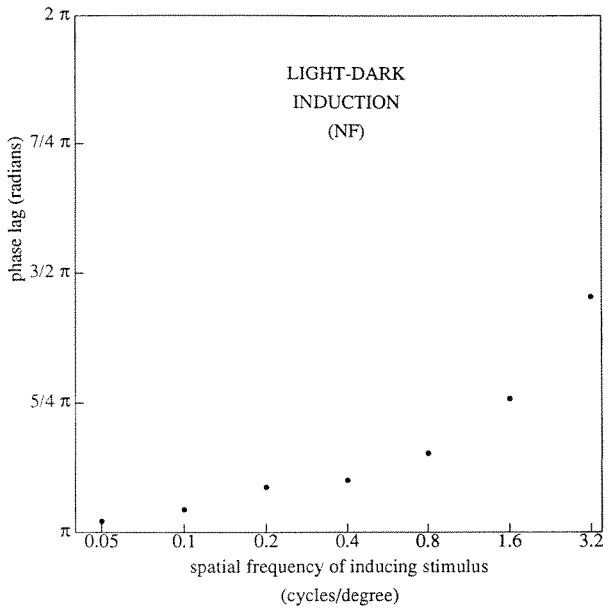


FIGURE 8. Predicted phase lag of induced temporal modulation with respect to the phase of the inner edge of the surround for observer NF.

After the parameters of the spatial weighting function had been estimated, they were used to derive predictions for $Y(\rho_0, \phi_i, \phi_j)$, the modulation induced by "light–dark" surrounds consisting of the pair of spatial frequencies (ϕ_i, ϕ_j). The expression for instantaneous induction [equation (2)] was expanded to equation (5):

$$y(t, \phi_i, \phi_j) = -\frac{A}{2}\int_L^X \kappa e^{-\alpha s}\{\cos[2\pi(\rho_0 t - \phi_i s + \phi_i L)] + \cos[2\pi(\rho_0 t - \phi_j s + \phi_j L)]\}\mathrm{d}s. \quad (5)$$

The expression for $Y(\rho_0, \phi_i, \phi_j)$ was derived from the temporal Fourier transform of equation (5). Because of the linear properties of the Fourier transform, the expression for $Y(\rho_0, \phi_i, \phi_j)$ was a straightforward generalization of the expression for $Y(\rho_0, \phi_i)$. Predictions for the nulling modulations required for compound stimuli are shown in Fig. 9. The predictions are generally consistent with the form of the data in Fig. 5(a), in terms of the shape of the curves, their parallelism, and the slight protuberances for the single frequency surrounds. The discrepancies between the two sets of curves are due mainly to two reasons. First, for this observer, the amount of nulling modulation required for the highest frequency surround was over-estimated by the model [Fig. 6(b)]. As a result, the curves in Fig. 9 have a shallower fall-off than the data and the lowest one is shifted upwards. The second reason is that this model does not predict phase shifts accurately, so that the sub-additivity of compound surrounds is less pronounced in Fig. 9 than in Fig. 5(a). In calculating phase delays, the present model assumes that the integrated effect of the surround on the test is instantaneous. It is more probable that signals require time to travel across lateral connections. Preliminary measurements indicate that, contrary to the model's predictions, the best nulling modulation for a spatially uniform surround has a small but significant phase delay with respect to the surround modulation. Precise measurements of the phase spectrum of the induced modulation are needed before any parameters are added to the spatial weighted-integration model. These estimates may also provide insights into the temporal properties of lateral interactions.

SPATIAL NONLINEARITY OF CHROMATIC INDUCTION

The chromatic induction results in Fig. 5(b, c) are clearly inconsistent with the assumptions of the weighted spatial integration model. The curves higher on the graph have steeper fall-offs, and the curves on the bottom are shallower. The non-additivity could conceivably be due to nonlinear spatial interactions and/or to point-wise nonlinearities. In this section point-wise nonlinearities will be ruled out as the cause.

For spatially uniform annular surrounds, Krauskopf *et al.* (1986) measured the amplitude of the nulling modulation as a function of the amplitude of the inducing stimulus in the three cardinal directions. This experiment revealed that whereas the nulling amplitude was a linear function of the inducing amplitude in the "light–dark" case, for chromatic induction, nulling amplitudes were compressive cubic polynomial functions of the inducing amplitudes.

When these estimates were used in numerical simulations, the results showed that adding a point-wise compressive nonlinearity to the model actually changed the predicted pattern of the results in the direction opposite to the nonadditivity in the empirical measurements. For the numerical simulation, the model for the integration of induced contrast along radial lines was elaborated to equation (6):

$$y(t, \phi_0) = \int_L^X \kappa e^{-\alpha s} f[A(s)]\mathrm{d}s \quad (6)$$

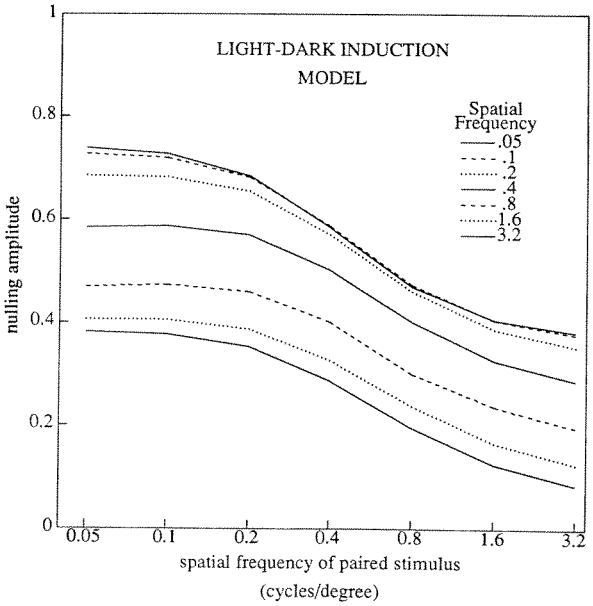


FIGURE 9. Predicted amplitude of "light–dark" nulling modulation for each of the compound inducing stimuli in Expt 2. Plotting conventions similar to Fig. 5(a), to which the predictions should be compared.

where

$$f[A(s)] = \beta A(s) - \gamma[A(s)]^3. \quad (7)$$

The curves in Fig. 10 were generated by numerical integration of the point-wise non-linear model. Representative values of β and γ from Krauskopf *et al.* (1986) were used in conjunction with the same values of κ and α as in Fig. 9. The curves represent the predicted amplitudes of nulling modulation at a temporal frequency equal to ρ_0. The predicted amplitudes at the third harmonic were around ten percent of these values. The differences between Figs 9 and 10 illustrate the effect of adding a compressive cubic polynomial. The right-most columns of points are identical in the two figures, whereas the left-most column is severely compressed in Fig. 10. The compressive nonlinearity disproportionately attenuates the induced effect of low-frequency surrounds. In Fig. 10, the departure from parallelism takes the form of a fanning out of the curves towards the right. This pattern of results will be predicted by every combination of a $g(s)$ which is a monotonically decreasing function of distance and an $f[A(s)]$ that is a compressive function of inducing amplitude. On the other hand, the chromatic induction curves [Fig. 5(b, c)] converge towards the right. Consequently, the non-additivity of chromatic induction is not due to a point-wise contrast compression, but is a result of nonlinear spatial interactions. Even though a regular temporal modulation at a higher harmonic was not noticeable inside the test at the null setting, systematic measurements of modulation induced at harmonics higher than the inducing modulation may provide an insight into these nonlinearities.

SUMMARY AND DISCUSSION

The results of the present study show that "light–dark" induction can be characterized as a linear spatial process, in the sense that the total induced effect of a circularly symmetric surround is equal to the sum of the induced effects of different parts of the surround. In addition, the effect of each part is weighted by a negative exponential function of the distance from the test. The magnitude of pure chromatic induction, however, is not an additive function of the surround effects. This nonlinearity is due to the nature of spatial interactions in the processing of chromatic signals, and is not due to a pointwise contrast nonlinearity. Both the theory and the experiments in this study deal with variations in spatial patterns at a punctate level. The present results provide information about the fine structure of the lateral interactions involved in color computations. In this aspect, this study goes beyond previous studies of the spatial weighting function (Yund & Armington, 1975; Krauskopf & Zaidi, 1986; Blackwell & Buchsbaum, 1988) and spatial additivity of induced contrast (Alpern & David, 1959; Zaidi & Krauskopf, 1987; Reid & Shapley, 1988; Shevell, Holliday & Whittle, 1988).

The "light–dark" results are interesting because they show supra-threshold spatial additivity across a fairly large region of visual space. Early stages of the visual system are composed of cells that have small receptive fields and fairly narrow bandpass sensitivity in the spatial frequency domain (for references see Shapley & Lennie, 1985). In addition, psychophysical results are consistent with a small number of classes of spatial frequency selective mechanisms functioning independently at threshold (for references see Graham, 1989). These mechanisms are generally thought to have spatial frequency bandwidths that are less than one octave wide. In "light–dark" induction, the spatial additivity covered spatial scales over a range of six octaves from 0.05 to 3.2 c/deg. Thus, at some stage of the visual system, important for the computation of brightness contrast, the outputs of these mechanisms must be summed in a fairly simple manner to lead to point-by-point additivity of lateral effects.

The induced effects of compound chromatic surrounds in Fig. 5(b, c) show that high frequency chromatic components severely attenuate the combined effect. This effect is particularly noticeable in the shallowness of the bottom curves in Fig. 5(b, c). Shevell and Wesner (1990) showed that the induced effect of a spatially uniform red annulus on a central test is more severely attenuated by a thin equiluminant white ring embedded in the surround than by a dark ring. If their results are analyzed in terms of spatial frequency, a thin white ring, equiluminant with the red surround, adds high frequency chromatic components, whereas a thin dark ring adds high frequency radiance components. The results of Shevell and Wesner's study show that adding high spatial frequency chromatic components to a uniform surround attenuates the induced effect considerably more than adding high spatial frequency radiance components, thus, corroborating one of the main results of the present study.

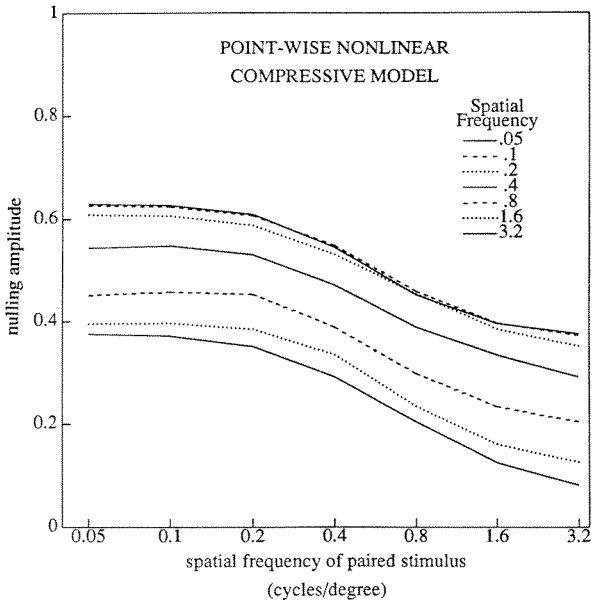


FIGURE 10. The effect of adding a point-wise compressive nonlinearity to the predictions in Fig. 9.

It should be emphasized that the present study considered only circularly symmetric surrounds. For "light–dark" induction, the results showed additivity of effects along lines radiating from the center of the test. A comprehensive test of the model in equation (1) requires surrounds that also vary tangential to the radial lines. When surrounds are not circularly symmetric, different colors are contiguous to different parts of a test region. There is some evidence that the induced effects due to different colors adjacent to the test region combine in a nonlinear fashion (Zaidi, 1990), but a comprehensive investigation of these spatial interactions remains to be done.

REFERENCES

Alpern, M. & David, H. (1959). The additivity of contrast in the human eye. *Journal of General Physiology, 43*, 109–126.

Blackwell, K. T. & Buchsbaum, G. (1988). The effect of spatial and chromatic parameters on chromatic induction. *Color Research and Application, 13*, 166–173.

Brindley, G. S. (1960). *Physiology of the retina and visual pathway*. Baltimore, Md: Williams & Wilkins.

Chevreul, M. E. (1839). *De la loi du contraste simultane des couleurs*. Paris: Pitois–Levreault.

Graham, N. (1989). *Visual pattern analyzers*. New York: Oxford University Press.

von Helmholz, H. (1924). *Physiological optics* (Vol. 2., pp. 264–271). Washington, D.C.: Optical Society of America.

MacLeod, D. I. A. & Boynton, R. M. (1978). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of the Optical Society of America, 69*, 1183–1186.

Krauskopf, J. & Zaidi, Q. (1986). Spatial factors in color induction. *Investigative Opthalmology and Visual Science, Suppl. 27*, 291.

Krauskopf, J., Williams, D. R. & Heely, D. M. (1982). The cardinal directions of color space. *Vision Research, 22*, 1123–1131.

Krauskopf, J., Zaidi, Q. & Mandler, M. B. (1986). Mechanisms of simultaneous color induction. *Journal of the Optical Society of America, A3*, 1752–1757.

Reid, R. C. & Shapley, R. (1988). Brightness induction by local contrast and the spatial dependence of assimilation. *Vision Research, 28*, 115–132.

Sachtler, W. L. & Zaidi, Q. (1991). Chromatic and luminance signals in visual memory. *Journal of the Optical Society of America. A9*.

Shapiro, A. & Zaidi, Q. (1991). The effect of prolonged temporal modulation on the response of color mechanisms. *Investigative Opthalmology and Visual Science, Suppl. 32*, 842.

Shapley, R. M. & Lennie, P. (1985). Spatial frequency analysis in the visual system. *Annual Reviews of Neuroscience, 8*, 547–583.

Shevell, S. K. & Wesner, M. F. (1990). Chromatic adapting effect of an achromatic light. *OSA Technical Digest Series, 15*, 149.

Shevell, S. K., Holliday, I. E. & Whittle, P. (1988). Brightness contrast from adjacent and nonadjacent lights. *OSA Technical Digest Series, 11*, 104.

Smith, V. C. & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 700 nm. *Vision Research, 15*, 161–171.

Walvaren, J. (1976). Discounting the background—the missing link in the explanation of chromatic induction. *Vision Research, 16*, 289–295.

Yund, E. W. & Armington, J. C. (1975). Color and brightness contrast effects as a function of spatial variables. *Vision Research, 15*, 917–929.

Zaidi, Q. (1990). Apparent brightness in complex displays: A reply to Moulden and Kingdom. *Vision Research, 30*, 1253–1255.

Zaidi, Q. & Halevy, D. (1992). Visual mechanisms that signal the direction of color changes. *Vision Research*. Submitted.

Zaidi, Q. & Krauskopf, J. (1987). Spatial interaction in color induction. *Investigative Opthalmology and Visual Science, Suppl. 28*, 214.

Zaidi, Q., Yoshimi, B. & Flannigan, J. (1991). The influence of shape and perimeter-length on induced color contrast. *Journal of the Optical Society of America, A8*, 1810–1817.

Acknowledgements—Sean Daly's help in the preparation of this manuscript is gratefully acknowledged. This research was supported partially by the National Eye Institute through grant EY07556 to Q. Zaidi. Portions of this work have been presented at the 1988 and 1990 annual meetings of OSA and at IBRO 1991.

APPENDIX

In this Appendix we provide details of the derivation of equation (3) from equation (2), plus the derivation of expressions for the amplitude and phase of the modulation induced by sinusoidal surrounds.

First the derivation of equation (3). Equation (2) is reproduced as equation (A1) below:

$$y(t, \phi_i) = A \int_L^X g(s) \cos(2\pi\rho_0 t - 2\pi\phi_i s + 2\pi\phi_i L)\, \mathrm{d}s \qquad \text{(A1)}$$

where y is the induced contrast at time t, A is the amplitude of the inducing stimuli, $g(s)$ is the spatial weighting function, ρ_0 represents the temporal frequency of the drifted sine wave, and ϕ_i is the spatial frequency of the drifted sine wave.

In order to estimate the amplitude and phase of the time varying components, we took the Fourier transform of both sides of the previous equation w.r.t. time which resulted in having to solve the following equation:

$$Y(\rho_0, \phi_i) = A \int_L^X g(s) \left[\int_{-\infty}^{\infty} \cos(2\pi\rho_0 t - 2\pi\phi_i s + 2\pi\phi_i L) e^{-i2\pi\rho_0 t} \mathrm{d}t \right] \mathrm{d}s. \qquad \text{(A2)}$$

The next step was to solve the integral inside the square brackets:

$$\int_{-\infty}^{\infty} \cos(2\pi\rho_0 t - 2\pi\phi_i s + 2\pi\phi_i L) e^{-i2\pi\rho_0 t}\, \mathrm{d}t \qquad \text{(A3)}$$

using the substitutions

$$v = \frac{\rho_0}{\phi_i};\ z = t - \frac{s}{v} + \frac{L}{v};\ \mathrm{d}z = \mathrm{d}t$$

the expression reduced to:

$$e^{-i2\pi\frac{\rho_0}{v}(s-L)} \int_{-\infty}^{\infty} \cos(2\pi\rho_0 z) e^{-i2\pi\rho_0 z}\, \mathrm{d}z. \qquad \text{(A4)}$$

Since the integral is the Fourier transform of a cosine function, the expression is equal to:

$$e^{-i2\pi\phi_i s} e^{i2\pi\phi_i L} \left[\frac{1}{2} [\delta(\rho - \rho_0) + \delta(\rho + \rho_0)] \right]. \qquad \text{(A5)}$$

Expression (A5) was substituted into equation (A2), giving a form identical to equation (3):

$$Y(\rho_0, \phi_i) = \frac{A}{2} \left[\delta(\rho - \rho_0) + \delta(\rho + \rho_0) \right] e^{i2\pi\phi_i L} \int_L^X g(s) e^{-i2\pi\phi_i s}\, \mathrm{d}s. \qquad \text{(A6)}$$

To derive expressions for the amplitude and phase of induced modulation, the exponential form of the spatial weighting function, given in equation (4), was substituted into equation (A6).

$$Y(\rho_0, \phi_i) = \frac{A}{2} \left[\delta(\rho - \rho_0) + \delta(\rho + \rho_0) \right] e^{i2\pi\phi_i L} \int_L^X \kappa e^{-\alpha s} e^{-i2\pi\phi_i s}\, \mathrm{d}s. \qquad \text{(A7)}$$

The solution of the integral on the right hand side gives the following expression:

$$\frac{\kappa}{-(\alpha + i2\pi\phi_i)}\left[e^{-(\alpha + i2\pi\phi_i)X} - e^{-(\alpha + i2\pi\phi_i)L}\right]. \quad \text{(A8)}$$

After substituting equation (A8) into equation (A7), the right hand side was converted into polar form to give the amplitude (A9) and the phase (A10) of the induced temporal modulation.

$$\text{Amp}[Y(\rho_0, \phi_i)] = A\frac{\kappa}{\sqrt{\alpha^2 + 4\pi^2\phi_i^2}} \times \sqrt{e^{-2\alpha X} + e^{-2\alpha L} - 2e^{-\alpha(X+L)}\cos[2\pi\phi_i(X - L)]} \quad \text{(A9)}$$

$$\text{Phase}[Y(\rho_0,\phi_i)] = \text{atan}\left[\frac{e^{-\alpha X}\sin\left(\text{atan}\left[\frac{2\pi\phi_i}{\alpha}\right] + 2\pi\phi_i X\right) - e^{-\alpha L}\sin\left(\text{atan}\left[\frac{2\pi\phi_i}{\alpha}\right] + 2\pi\phi_i L\right)}{e^{-\alpha X}\cos\left(\text{atan}\left[\frac{2\pi\phi_i}{\alpha}\right] + 2\pi\phi_i X\right) - e^{-\alpha L}\cos\left(\text{atan}\left[\frac{2\pi\phi_i}{\alpha}\right] + 2\pi\phi_i L\right)}\right]. \quad \text{(A10)}$$

Equation (A9) was fit to the data in Fig. 6 to estimate κ and α. The predicted phase lags are plotted in Fig. 8.