A Role for Cortical Crosstalk in the Binding Problem: Stimulus-driven Correlations that Link Color, Form, and Motion

Vincent A. Billock[1] and Brian H. Tsou[2]

Abstract

■ The putative independence of cortical mechanisms for color, form, and motion raises the binding problem—how is neural activity coordinated to create unified and correctly segmented percepts? Binding could be guided by stimulus-driven correlations between mechanisms, but the nature of these correlations is largely unexplored and no one has (intentionally) studied effects on binding if this joint information is compromised. Here, we develop a theoretical framework which: (1) describes crosstalk-generated correlations between cortical mechanisms for color, achromatic form, and motion, which arise from retinogeniculate encoding; (2) shows how these correlations can facilitate synchronization, segmentation, and binding; (3) provides a basis for understanding perceptual oddities and binding failures that occur for equiluminant and stabilized images. These ideas can be tested by measuring both perceptual events and neural activity while achromatic border contrast or stabilized image velocity is manipulated. ■

INTRODUCTION

How does the brain coordinate the activity of visual mechanisms to create unified percepts? There are two particularly important aspects to this binding problem (see Treisman, 1996 for a review of all aspects)—the binding of different parts of the same image and the binding of different image features (e.g., form, color, and motion). Part- and feature-binding are closely related because each feature class contributes to the segmentation of images from backgrounds (Regan, 2000). This article is not a theory of binding, nor is it a theory of motion, color, and luminance defined form—on which vast literatures already exist. Rather, we create a framework in which binding and its perceptual attributes may be better understood. Towards this end, we tackle some key questions that are too seldom addressed in binding studies: (1) Just how independent are putatively independent cortical mechanisms for color, achromatic form, and motion, and what is the nature of any crosstalk between them? (2) How could crosstalk between these mechanisms contribute to binding? (3) If ordinary vision is a resounding binding success, what would a catastrophic part-binding failure look like (feature-binding failures are called illusory conjunctions; Treisman, 1996), and under what conditions would binding failures occur?

Background on Parallel Perceptual Mechanisms and Binding Theory

Ample physiological and psychophysical evidence suggest that human vision is mediated by multiple mechanisms, each responding to selective combinations of image attributes (color, spatial frequency, orientation, motion, etc.; for reviews see Regan, 2000; Zeki, 1993). This parallel processing—although crucial to understanding visual detection and appearance—raises a binding problem: How are unified and correctly segmented percepts created from the activity of putatively independent, spatially separated cortical mechanisms? One theory is that binding stems from coordination of mechanisms responding to an image (von der Malsburg, 1981, 1995). Electrophysiological studies show coordinated γ-band (generally 30–90 Hz) oscillations between cells in the same and different orientation columns of the same cortical area, in different cortical areas, and in different hemispheres (for review and debates, see Gray, 1999; Shadlen & Movshon, 1999; Singer, 1999; Singer & Gray, 1995). Recently, a series of compelling studies have tied specific γ-band EEG activity in humans to specific percepts (Gruber, Müller, & Keil, 2002; Tallon-Baudry & Bertrand, 1999). This impressive corpus of work on binding and neural synchronization compels us to address a key question: What information governs binding?

In principle, a sufficiently potent coupling could synchronize any two neurons' activity, but indiscrimi-

[1]General Dynamics, [2]U.S. Air Force Research Laboratory

nant perceptual binding is undesirable. There are two key kinds of information available to bind sensory mechanisms: location coding and correlated activity. Some studies suggest that binding exploits location coding; for example, two cells in different cortical areas, each responding to a different image feature (e.g., color and orientation), are likely to synchronize if both are active and both are coded for the same retinotopic location (for review, see Finkel & Edelman, 1989). This is plausible because of the substantial retinotopically specific connections between cortical areas. We do not denigrate the role of retinotopic mapping in binding, but because large receptive fields in the extrastriate cortex degrade localization, it is useful to have an independent supplemental source of information for binding. Moreover, supplementing location coding should speed and improve binding (time is at a premium when segmentation/binding takes only 50–200 msec; von der Malsburg, 1999). We suggest that a type of stimulus-specific common neural input may provide information useful for binding.

THEORETICAL BUILDING BLOCKS

A Role for Stimulus-induced Correlations in Oscillatory Binding

A tenet of binding theory is that binding is driven by (and makes manifest) correlations between mechanisms responding to common image features. In principle, an initial weak correlation between oscillatory mechanisms is amplified into mutual synchronization (von der Malsburg, 1981). However, what guarantees that synchronization will be correctly linked to real stimulus features? To avoid binding everything, competitive networks could select the strongest correlations to be synchronized. Correlated activity in neural mechanisms is induced by causal relations between temporal events (e.g., by perception–action interactions) and by common neural inputs (Feng, 2000; von der Malsburg, 1999). One particularly promising source of correlated activity—crosstalk—is both stimulus-driven and induced by common neural inputs. Crosstalk is the unintended portion of a selectively filtered signal. For example, P cells carry both hue and luminance information (De Valois & De Valois, 1988). A cortical cell that filters the P-cell input to extract luminance may also extract a portion of the chromatic signal as well (Billock, 1995). This crosstalk—a nuisance in visual parallel processing theory—has been neglected in binding theory. The closest exception is Horn, Sagi, and Usher's (1991) segmentation/binding network, which uses correlated noise to synchronize neural oscillators responding to different visual modalities. Consider two such systems (E_1, E_2) coupled through a common set of inhibitory (I) neurons.

$$
\begin{aligned}
\mathrm{d}E_1/\mathrm{d}t &= -E_1 + F_\tau \\
&\qquad (aE_1 - bI - \theta_\mathrm{E} - cR_1 + \mathrm{INPUT}_1 + Q_{1,2}) \\
\mathrm{d}R_1/\mathrm{d}t &= (1/d - 1)R_1 + E_1 \\
\mathrm{d}E_2/\mathrm{d}t &= -E_2 + F_\tau \\
&\qquad (aE_2 - bI - \theta_\mathrm{E} - cR_2 + \mathrm{INPUT}_2 + Q_{1,2}) \\
\mathrm{d}R_2/\mathrm{d}t &= (1/d - 1)R_2 + E_2 \\
\mathrm{d}I/\mathrm{d}t &= -I + F(fE_1 + gE_2 - bI - \theta_I) \qquad (1)
\end{aligned}
$$

Here E_1, E_2 are firing rates driven by different visual features (INPUT_1, INPUT_2), which in this article could be achromatic contrast, color, or motion. R_1, R_2 are dynamic thresholds; a, b, c, d, g, θ, τ are constants; and $F_\tau(x) = 1/(1+e^{-x/\tau})$. Key attributes of Equation 1 are that E_1 and E_2 oscillate (fire) in response to their sensory inputs and can synchronize those oscillations. The $Q_{1,2}$ term is an input common to both systems which speeds and strengthens synchronization—a result likely generic to coupled systems. In Horn et al.'s (1991) model, it represents correlated noise. Because Horn et al. consider binding of color and form, they speculate that some noisy ''early mixed representation of shape and color information exists in the input layer.'' Another model uses a cell class tuned along two stimulus dimensions and couples it to a cell class that has a common stimulus dimension (Roelfsema, Engel, König, & Singer, 1996; e.g., oriented cells for chromatic edges coupled to similarly oriented cells for achromatic contrast). The correlated noise and common dimension approaches are more similar than they seem. An ''early mixed representation'' is an apt description of retinogeniculate multiplexing. Attempts to create cortical cells tuned along (and labeled for) one stimulus dimension inevitably induce stimulus-specific crosstalk from other stimulus dimensions (e.g., Billock, 1995). Below, we quantify this crosstalk for two specific sensory interactions.

Fragmentation and Fading—Characteristic Failure Modes of Segmentation Networks

Generally, within a sensory modality, segmentation involves cooperation between mechanisms that agree and competition between mechanisms that disagree. These methods are generic and apply whether an image is to be segmented on the basis of achromatic form, color, motion, depth, and so forth (Levine, 2000; Wang & Terman, 1997; Grossberg & Wyse, 1992; von der Malsberg & Buhmann, 1992). Competitive networks define the location of edges and borders. Cooperative networks fill in areas between borders and reinforce the response of retinotopically neighboring cells signaling

an edge. Such segmentation systems have characteristic failure modes: If cooperation runs amok, then all mechanisms agree and the image is grayed-out, rather than segmented. If such a failure occurs in an already segmented image, the image would appear to melt or fade. If competition is too strong, or cooperation too weak, then the image may fragment. For example, Terman and Wang's (1995) neural network (using nearest neighbor connectivity) tends to fragment images. Similarly, stereopsis segmentation networks fragment images if thresholds for keeping cells active and/or entraining adjacent inactive cells are set too high (Marr, Palm, & Poggio, 1978). Below, we examine situations (equiluminance and retinal stabilization) where percepts fade, melt, and/or fragment and argue these phenomena are understandable in a segmentation/part-binding framework. These fading/fragmentation problems may be alleviated if multiple segmentation networks share information about location of borders. Reinforcement between parallel networks—called ''cross-modal construction'' (Finkel & Edelman, 1989)—has the potential to reinforce and refine estimates of image features (Roelfsema et al., 1996; Schillen & König, 1994; Horn et al., 1991; Poggio, Gamble, & Little, 1988). Cross-modal construction is intuitive: Realistic segmentation networks have response thresholds; weak cooperative interactions between segmentation networks responding to different features of the same image (e.g., chromatic and achromatic form) allow the more active network to give the other network's units a boost above threshold, but do not entrain units not driven by sensory input, if their thresholds are chosen appropriately (Horn et al., 1991). We argue that cross-modal construction depends on a stimulus-linked variant of the correlated noise that Horn used for synchronization.

ORIGIN AND NATURE OF CROSSTALK-BASED CORRELATIONS BETWEEN ACHROMATIC AND CHROMATIC CORTICAL FORM MECHANISMS

Retinogeniculate Origins of Correlations between Chromatic and Achromatic Mechanisms

There is much evidence for chromatic/achromatic interactions in human vision (Mullen & Kingdom, 1991). Some interactions arise early in the visual pathway. About 80% of LGN cells have Type I receptive fields (center of different spectral sensitivity than its spatially opponent surround; Wiesel & Hubel, 1966); these P cells respond to both chromatic and achromatic stimuli (De Valois & Pease, 1971). P cells account for both the threshold chromatic and achromatic spatio-temporal contrast sensitivity functions (except for the achromatic low spatial, high temporal frequency corner; Merigan & Maunsell, 1993; Kelly, 1983). Although the P cell's mixed signal seems ambiguous, Ingling and Martinez's (1983, 1985) algebraic identity factors the P cell's sensitivity into psychophysically meaningful terms. The expansion for $r{+}g{-}$ cells (similar identities obtain for other cell types and for mixed cone surrounds; Billock, 1996) is:

$$\begin{aligned} P_{r+g-} &= RS_eT_e - GS_iT_i \\ &= (R+G)\{(S_e - S_i)(T_e + S_i) + (S_e + S_i)(T_e - T_i)\}/4 \\ &\quad \text{(Achromatic, Spatio-temporally bandpass)} \\ &\quad + (R-G)\{(S_e - S_i)(T_e - T_i) + (S_e + S_i)(T_e + T_i)\}/4 \\ &\quad \text{(Chromatic, Spatio-temporally lowpass)} \end{aligned} \tag{2}$$

Here, R and G are absorptions in L- and M-cones, S_e, S_i are center (excitatory) and surround (inhibitory) spatial weighting functions, and T_e, T_i are center and surround temporal impulse response functions. S_e, S_i, T_e, T_i are all lowpass functions of spatial/temporal frequency, so their sums are lowpass and differences are bandpass. A variation of Equation 2 models both the chromatic and achromatic contrast sensitivity functions (Burbeck & Kelly, 1980). For our current purposes, this formulation is unnecessarily complex; we develop some basic ideas using a simplification and then return to the full version when necessary. Temporarily neglecting time gives

$$\begin{aligned} P_{r+g-} &= R_{\text{Center}} - G_{\text{Surround}} \\ &= (R+G)(S_e - S_i)/2 \\ &\quad \text{(achromatic, bandpass tuning)} \\ &\quad +(R-G)(S_e+S_i)/2 \\ &\quad \text{(chromatic, lowpass tuning)} \end{aligned} \tag{3}$$

For simplicity, we restrict space to one dimension (x) and define the stimulus in terms of L and M cone absorptions: $Z(x) = \{R(x), G(x)\}$. The response of the P cell to the stimulus is

$$\begin{aligned} P_{r+g-}(x) \otimes Z(x) &= 0.5\{R(x) + G(x)\} \otimes \{S_e(x) - S_i(x)\} \\ &\quad \text{(Achromatic response)} \\ &\quad + 0.5\{R(x) - G(x)\} \otimes \{S_e(x) + S_i(x)\} \\ &\quad \text{(Chromatic response)} \end{aligned} \tag{4}$$

where the convolution of two functions $A(x) \otimes B(x) = \int A(\tau)B(x - \tau)d\tau$ and τ is a dummy variable of integration. Equations 3 and 4 reveal a subtle encoding of chromatic and achromatic information that is decipherable by cortical decoders. To see how, let the excitatory center (S_e) and inhibitory surround (S_i) spatial weighting functions of this Type I cell be represented by Gaussians. The achromatic term is a difference of Gaussians that closely approximates a second spatial derivative of a Gaussian. The chromatic term is the sum of these Gaussians and is fit by a Gaussian with a space constant (σ) about 1.83 times that of the achromatic term. If some simple assumptions (Billock, 1995) hold,

then the spatial tuning of the P cell is the Fourier transform ($F[P]$) of Equation 3.

$$F[P_{r+g-}] = 0.5(R+G)a^2 f_s^2 H_\sigma(f_s) \quad \text{(Achromatic tuning)} + (R-G)H_{1.83\sigma}(f_s) \quad \text{(Chromatic tuning)} \tag{5}$$

where $H_\sigma(f_s)$ is $\exp[-2(\pi\sigma f_s)^2]$ (the Fourier transform of the center Gaussian), $a = 2\pi$, and f_s is spatial frequency. The achromatic term is a spatial bandpass filter; the chromatic term is lowpass, multiplexing the chromatic and achromatic information into different frequency ranges. If the multiplexing filters had no overlap, two well-chosen cortical filters could perfectly separate (demultiplex) the chromatic and achromatic information (Kingdom & Mullen, 1995; Billock, 1991, 1995). For P cells, the encoding filters overlap, inducing crosstalk in the cortical decoders, especially at moderate spatial frequencies where the overlap is greatest. Consider the extraction of achromatic information from P cells by the simplest spatial bandpass filter—spatial differentiation (Billock, 1995). Taking the nth local derivative of the P-cell array is equivalent to differentiating Equations 3 or 5, with respect to a spatial dimension; in neural terms it corresponds to building cortical dual-opponent cells (Thorell, De Valois, & Albrecht, 1984) by lateral inhibition between afferent geniculate cells (for details, see Billock, 1995). In the spatial domain, such derivative operators look like simple cells; 1-D differentiation produces oriented cells and the order of differentiation determines the number of alternating excitatory and inhibitory lobes; Equation 6 describes a cell that has $n + 3$ receptive field lobes when probed with achromatic spots, but $n + 1$ lobes for equiluminant hue probes. The spatial frequency response ($A(f_s)$) of such an achromatic mechanism to a stimulus (defined in terms of L- and M-cone absorptions) $Z(f_s) = \{R(f_s), G(f_s)\}$ is

$$\begin{aligned} A(f_s) &= Z(f_s)(F[D^n P(f_s)]) \\ &= \mathbf{0.5[}\boldsymbol{R(f_s)+G(f_s)]a^{n+2}f_s^{n+2}H_\sigma(f_s)} \\ &\quad \textbf{(achromatic response)} \\ &\quad + [R(f_s) - G(f_s)]a^n f_s^n H_{1.83\sigma}(f_S) \\ &\quad \text{(chromatic crosstalk)} \end{aligned} \tag{6}$$

This bandpass filtering model correctly predicts the spatial tuning of cortical cells sensitive to both achromatic and chromatic contrast (but presumably labeled only for achromatic contrast; Billock, 1995). Increasing n shifts the tuning to higher spatial frequencies and narrower bandwidths. The typical cortical cell has an achromatic bandwidth of 1.4 octaves (De Valois & De Valois, 1988) matching the achromatic term of Equation 6 if $n = 2$ (i.e., matched filtering for luminance signals in P cells). Similarly, color signals can be extracted by matched lowpass filtering (second-order integration), resulting in achromatic crosstalk at all but the lowest spatial frequencies.

$$\begin{aligned} C(f_S) &= Z(f_s)(F[D^{-2}P(f_s)]) \\ &= 0.5[R(f_s) + G(f_s)]H_{1.83\sigma}(f_s) \\ &\quad \text{(luminance crosstalk)} \\ &\quad + \mathbf{[}\boldsymbol{R(f_s)-G(f_s)]H_{3.35\sigma}(f_s)} \\ &\quad \textbf{(chromatic response)} \end{aligned} \tag{7}$$

An engineer would find this odd—this kind of crosstalk is generally undesired and avoided in communications systems. Yet, as discussed below, crosstalk is useful information for binding.

Estimation of Stimulus-induced Correlations between Spatial and Chromatic Mechanisms

The discussion of crosstalk above was deliberately oversimplified to illustrate general principles. To estimate crosstalk-based correlations, we generalize Equations 6 and 7, to include the temporal response. Selective two-dimensional (x, t) spatio-temporal matched filtering of Equation 2 yields

$$\begin{aligned} A(f_s,f_t) &= \mathbf{[}\boldsymbol{R(f_s,f_t)+G(f_s,f_t)][S_e(f_s)T_e(f_t)-S_i(f_s)T_i(f_t)]^2} \\ &\quad \textbf{\{Achromatic response\}} \\ &\quad + [R(f_s,f_t) - G(f_s,f_t)][S_e(f_s)T_e(f_t) \\ &\quad + S_i(f_s)T_i(f_t)][S_e(f_s)T_e(f_t) - S_i(f_s)T_i(f_t)] \\ &\quad \{\text{Chromatic Crosstalk}\} \end{aligned} \tag{8}$$

$$\begin{aligned} C(f_x,f_t) &= \mathbf{[}\boldsymbol{R(f_s,f_t)-G(f_s,f_t)][S_e(f_s)T_e(f_t)+S_i(f_s)T_i(f_t)]^2} \\ &\quad \textbf{\{Chromatic Response\}} \\ &\quad + [R(f_s,f_t) + G(f_s,f_t)][S_e(f_s)T_e(f_t) \\ &\quad + S_i(f_s)T_i(f_t)][S_e(f_s)T_e(f_t) - S_i(f_s)T_i(f_t)] \\ &\quad \{\text{Achromatic Crosstalk}\} \end{aligned} \tag{9}$$

Although not as elegant as Equations 6 and 7, Equations 8 and 9 are computable if the S_e, S_i, T_e, T_i functions can be estimated (see Kelly, 1989; Burbeck & Kelly, 1980; for methods and estimated functions). We could use Equations 8 and 9 as INPUT_1, INPUT_2 in Equation 1 and ignore $Q_{1,2}$; the correlated input is implicit. However, it is enlightening to use only the achromatic and chromatic response terms (the terms in bold type) as INPUT_1, INPUT_2 and to use Equations 8 and 9 (with the crosstalk terms) to estimate the stimulus-driven correlations ($Q_{1,2}$ in Equation 1) between spatial mechanisms that extract information from P cells about achromatic form and other mechanisms that extract information about color or chromatic form. We define spectral correlation as the integrated overlap of the Fourier spectra of two functions (an analog of Signal Detection

Theory's cross-ambiguity function). The spectral correlation between $A(f_s,f_t)$ and $C(f_s,f_t)$ is

$$Q_{1,2} = A(f_s,f_t) \circledcirc C(f_s,f_t) = \int A(\alpha,\beta)C(f_s+\alpha,f_t+\beta)d\alpha d\beta \qquad (10)$$

where α, β are dummy variables of integration. For an equiluminant stimulus, the achromatic signal to the achromatic mechanism is zero and the correlated crosstalk consists of the correlation between the chromatic signal in the chromatic pathway [the bold term in the $C(f_s)$ equation] and the chromatic crosstalk extracted by the achromatic pathway [the plain type term in the $A(f_s)$ equation]. Thus, at equiluminance, the visual system attempts to bind something to nothing, based on the correlation between something and bandpass filtered noise. This situation corresponds to some very odd perceptual effects.

Effects of Equiluminance on Perception

At equiluminance perception of form, depth and motion are degraded (Cavanagh, 1991; Livingstone & Hubel, 1987; Gregory, 1977). Equiluminance can also disrupt segmentation and binding. Segmentation by binocular disparity is severely degraded at equiluminance for random dot stereograms (Lu & Fender, 1972). Surfaces are not linked together if their features are defined only by color; an image made of equiluminant colors appears as patches of those colors, not as a unified whole (Cavanagh, 1991; Livingstone & Hubel, 1987). Moreover, perception of equiluminous forms can be unstable. Gregory (1977) found equiluminant images ''looked unstable in contrast and 'jazzy'.'' This instability is sometimes attributed to an inability of ocular accommodation to use color information, but this is contradicted by Kotulak, Morse, and Billock (1995). Liebmann (1927) found that there is a critical luminance zone within which ''everything flows ... glimmers ... everything is soft, jelly-like, colloidal. Often ... parts which belong together in the normal figure now have nothing to do with each other. (It is) a world without firm things, without solidity'' (translated in Cavanagh, 1991). Studies using luminance minimized borders show these effects are particularly severe for borders defined by S-cone (tritan) modulation; the borders collapse and the color fields blend into a continuous color gradient (Buck, Frome, & Boynton, 1977).

No Evidence for Explanations Based on Lack of Parvo Inputs to Central Pathways

Livingstone and Hubel (1987) ascribe the detrimental effects of equiluminance to a lack of parvo inputs to motion, form, and depth channels, but evidence contradicts this (Merigan & Maunsell, 1993; Ingling & Grigsby, 1990; Schiller, Logothesis, & Charles, 1990). Moreover, there are chromatic mechanisms sensitive to form, motion, and depth (Regan, 2000; Cavanagh, 1991; Mullen & Kingdom, 1991). Indeed, paradoxically, some chromatic phenomena are adversely affected by equiluminance, including color discrimination, color contingent aftereffects, and color rivalries (for review, see Mullen & Kingdom, 1991; Livingstone & Hubel, 1987), suggesting a more subtle origin for the effects of equiluminance. Several possibilities are considered below.

Evidence Contradicting Explanations Based on Low Chromatic Acuity

Although the low acuity of chromatic mechanisms affects some visual phenomena (Cavanagh, 1991), it does not explain the poor and unstable contrast of some equiluminous borders. Border sharpness is not strictly a high spatial frequency phenomenon; a blurred edge looks sharp if the missing harmonics of its Fourier series are below detection threshold (Campbell, Hopwell, & Johnstone, 1978). Similarly, reducing acuity by dimming illumination has little effect on contrast over a large range and edges modulated at 15 Hz appear sharp, even though acuity is reduced by a factor of 2.5 (Livingstone & Hubel, 1987). Finally, as Mullen and Kingdom (1991) put it, ''it would be surprising if the lower border distinctness rated for S cone mechanisms compared to M–L cone ones was due to their differences in acuity since a greater difference in acuity occurs between luminance and M–L chromatic mechanisms with no loss of border distinctness.''

Luminance as a Master Signal?

Gregory (1977) posits that luminance is a master signal necessary for demarcating borders. This fits luminance captures color phenomena but does not explain why tritan equiluminous borders suffer excessively relative to equiluminous borders that stimulate the red/green system. Nor can the vulnerability of tritanopic borders be due to a lack of S-cone driven color contrast mechanisms; double opponent b–y cells (Livingstone & Hubel, 1984), are well suited for transducing S-cone modulated chromatic contrast and multiple S-cone driven spatial frequency channels are found psychophysically (Humanski & Wilson, 1993).

A Binding-level Explanation?

A possible explanation for the insalience of tritanopic borders stems from the origins of the $r-g$, $r+g$ (luminance) and b–y signals. The major input of $r-g$

signal to the cortex comes from spatially opponent P cells that multiplex luminance and color signals, inducing crosstalk between cortical luminance and $r-g$ color mechanisms. Conversely, it appears that a major source of retinogeniculate y–b signal is from cells with spatially nonopponent Type II receptive fields (de Monasterio & Gouras, 1975; Wiesel & Hubel, 1966). These cells carry no multiplexed luminance signal and therefore do not induce correlations between cortical luminance and color detectors. The paucity of S-cone inputs to luminance (for review, see Cavanagh, 1991) removes another source of correlations between y–b and luminance mechanisms. If correlations between mechanisms responding to borders are exploited for binding, perhaps the y–b system, deprived of correlations to luminance borders, evolved only a weak input to the border contrast system. We suggest that binding allows achromatic mechanisms to reinforce $r-g$ mechanisms involved in segmentation and border formation (the luminance-driven form mechanism may need less reinforcement from the chromatic system because—as discussed later—it receives reinforcement from motion mechanisms for unstabilized stimuli).

CORRELATED CORTICAL SPATIAL/TEMPORAL MECHANISMS

Why Form and Motion (or Space and Time) Are Not Independent

There is evidence for separate mechanisms mediating detection of spatial and temporal variation (for review, see Zeki, 1993). There is also evidence for interactions between spatially and temporally tuned mechanisms (Burt, 1987; Breitmeyer & Ganz, 1976). Such interactions—as unavoidable as chromatic/achromatic interactions—are due to the physical entanglement of spatial and temporal information, and to spatio-temporal multiplexing in retinogeniculate neurons. The physical entanglement of motion and form is obvious; the velocity (deg/sec) of a moving grating is $V = f_t$ (cycles/sec)/f_s (cycles/deg), where f_s, f_t are spatial and temporal frequency (each spatial and temporal frequency in the stimulus's Fourier representation can be so treated). The neural entanglement of spatial and temporal information is more complicated, but can be analyzed analogously to chromatic/achromatic interactions. Consider the psychophysical achromatic spatio-temporal contrast sensitivity function. In principle, if the function is separable (decomposable into the product of spatial and temporal functions), then independent spatial and temporal information can be extracted by a homomorphic filter (a log transform followed by a matched filter). Unfortunately, contrast sensitivity is inseparable. However, Burbeck and Kelly (1980) showed that an Ingling–Martinez identity could achieve a limited separation, decomposing contrast sensitivity into the sum of two separable surfaces. For a stimulus $Z(f_x,f_t)$ the frequency response (assuming equal integrated excitatory/inhibitory sensitivity) is

$$X(f_x,f_t) = 0.5Z(f_x,f_t)[S_e(f_x) - S_i(f_x)][T_e(f_t) + T_i(f_t)]$$
Sustained (spatially tuned) response
$$+ 0.5Z(f_x,f_t)[S_e(f_x) + S_i(f_x)][T_e(f_t) - T_i(f_t)]$$
Transient (temporally tuned) response
(11)

where S_e, T_e, S_i, T_i are the excitatory and inhibitory spatial and temporal response functions defined in Equation 2. In Equation 11, the terms are labeled sustained and transient. The sustained term has bandpass tuning for spatial frequencies and lowpass tuning for temporal frequencies. The transient term has lowpass tuning for spatial frequency and bandpass tuning for temporal frequency. In the psychophysical literature, the terms sustained and transient usually refer to separate retinogeniculate spatial and temporal processing pathways. Yet, Burbeck and Kelly's analysis shows that both kinds of responses are embedded in all retinogeniculate cells with center/surround receptive fields, and are well modeled by Equation 11 for cells that obey superposition (P cells and X-like M cells). Moreover, although the sustained and transient signals can be separated algebraically, there is no plausible spatio-temporal filter that enables a strict spatial and temporal separation by physiological means; the spectral content of the terms overlap and attempts to extract one signal will extract a small crosstalk signal as well.

Quantifying Stimulus-induced Spatial and Temporal Mechanism Correlations

Here, we estimate the stimulus-driven correlation between temporally and spatially tuned pathways, created by cortical filtering of the spatio-temporal signals carried by LGN afferents. To extract the maximum signal from the spatially tuned (sustained) component of an array of active X-cells, we apply a filter matched to the sustained component, yielding a frequency response

$$\mathrm{SP}(f_x,f_t) = \mathbf{0.5Z(f_x,f_t)[S_e(f_x) - S_i(f_x)]^2[T_e(f_t) + T_i(f_t)]^2}$$
(Spatially tuned term)
$$+ 0.5Z(f_x,f_t)[S_e(f_x)^2 - S_i(f_x)^2][T_e(f_t)^2 - T_i(f_t)^2]$$
(Common information term)
(12)

Similarly, a matched filter can be used to attempt to extract the temporally tuned (transient) component of the X-cell signals, yielding a frequency response

$$\mathrm{TE}(f_x,f_t) = 0.5Z(f_x,f_t)[S_e(f_x)^2 - S_i(f_x)^2][T_e(f_t)^2 - T_i(f_t)^2]$$
(common information term)
$$+ \mathbf{0.5Z(f_x,f_t)[S_e(f_x) + S_i(f_x)]^2[T_e(f_t) - T_i(f_t)]^2}$$
(temporally tuned term)
(13)

Note that the matched filters are only partially successful in extracting signals tuned along one stimulus dimension. In both cases, the undesired portions of the signal—tuned along both space and time—are identical, and so spectral correlation between the mechanisms is dominated by the crosstalk. As before, we use the bold terms of Equations 12 and 13 as input signals INPUT_1, INPUT_2 in Equation 1 and estimate correlated crosstalk ($Q_{1,2}$ in Equation 1) by computing the spectral correlation between Equations 12 and 13 (neglecting unknown sources of correlated noise; e.g., Horn et al., 1991)

$$Q_{1,2} = \text{SP}(f_s,f_t) \circledcirc \text{TE}(f_s,f_t) = \iint \text{SP}(\alpha,\beta)\text{TE}(f_s+\alpha,f_t+b)d\alpha d\beta \qquad (14)$$

where α and β are dummy variables of integration. These spatio-temporal interactions present two possibilities for crosstalk-mediated segmentation: (1) Mutual feedback between correlated form and motion mechanisms both responding to moving edges (the cross-modal construction model discussed above for form/color interactions); (2) Motion-correlated segmentation within the form pathway; that is, stimulus motion gives rise to two responses in the form pathway: the sustained form signal (the first term in Equation 12) and the transient crosstalk signal (the second term in Equation 12). Because these crosstalk signals arise only in form detectors stimulated by the target's motion, they are a potential cue for binding by correlated motion.

A Binding Failure—Perception of Stabilized Images

During normal vision images are in constant motion; microtremors of the eye make volitional stabilization almost impossible. However, images can be stabilized mechanically or by producing afterimages on the retina. These stabilized images can fade away rather quickly, an effect often attributed to the temporal response properties of retinogeniculate neurons. We argue that this explanation is grossly inadequate and that a failure of cortical segmentation mechanisms is indicated. However, the cortex can only operate on what afferent mechanisms send it, so next we analyze what early and cortical mechanisms contribute to stabilized image perception.

What Early Mechanisms Do to Perception of Stabilized Images

Some stabilized image percepts are consistent with retinogeniculate cell properties. Low contrast, low spatial frequency stimuli (favoring transduction by transient magno [M] cells) usually fade faster than high contrast, high spatial frequency stimuli (favoring transduction by sustained parvo [P] cells; Ingling & Grigsby, 1990). Moreover, nonbleaching chromatic afterimages elevate both achromatic and chromatic thresholds, suggesting a common P cell-like pathway (Kelly & Martinez-Uriegas, 1993). There are some sustained M cells that respond to static images for a few seconds, but their temporal response becomes transient for high contrast stimuli (Benardete, Kaplan, & Knight, 1992), leaving only P cells to mediate perception of afterimages. Moreover, King-Smith, Rosten, and Alvarez (1980) describe a subject who (on psychophysical grounds) appears to be missing the parvo system; this subject was also unable to perceive afterimages. Hence, transduction by P cells is a necessary condition for perception of most stabilized images. However, retinogeniculate properties cannot account for at least seven lines of evidence that central mechanisms are responsible for some binding-failure-like oddities of stabilized images.

Three Properties of Stabilized Image Perception Paradoxical to P- and M-Cell Properties

(1) Anomalously Rapid Fading of Images that Should Tap High Acuity Sustained Channels

Recall that Equation 2 shows that P cells have a sustained response for high spatial frequencies. Near the fovea, these cells have midget receptive fields capable of transducing spatial frequencies up to 60 c/deg (cone sampling limit). Additionally, these P cells (unlike M cells) are sensitive to high contrasts (Benardete et al., 1992). So, it is odd that the stabilized pattern of high contrast shadows cast by blood vessels on the retina fades faster than other stabilized images (Coppola & Purves, 1996). Moreover, the closer the blood vessels are to the fovea, the faster they fade (as fast as 80 msec for the highest spatial frequency components), in contradiction to the response characteristics and retinal distribution of P cells.

(2) Anomalously Rapid Fading for Stimuli that Tap Chromatic Sustained Mechanisms

P cells have a chromatic response that is lowpass in both space and time (Equation 2 shows a bandpass chromatic component as well, but this merely adds to the response at moderate spatial and temporal frequencies without driving down the response to low frequencies). Although Equation 2 must hold if P cells obey linear superposition, it does not model psychophysical chromatic contrast sensitivity for very low temporal frequencies (<0.2 Hz), where the temporal CSF slope is consistent with a first-order temporal differentiation (Kelly, 1981); this extra temporal derivative is probably a cortical process acting on P-cell inputs. A perceptual manifestation of this is the dramatic elevation of stabilized chromatic grating detection thresholds (a factor of at least 45 greater than for chromatic unstabilized gratings and much higher than the elevation for achromatic

gratings; Kelly, 1983). To reemphasize, unlike achromatic information, which is also transmitted by magno units, color is carried by P cells with sustained temporal properties. If retinogeniculate mechanisms are solely responsible for fading of stabilized images, one would expect color to be affected less by stabilization relative to stabilized luminance signals, not more.

(3) Fading of Dynamic Images is Paradoxical to a Peripheral Mechanism Explanation

Empty figures defined by a twinkling random dot background fade if fixated steadily (Spillman & Kurtenbach, 1992). Dynamic image fading cannot be due to transient peripheral mechanisms; fading is actually faster for dynamic stimuli than for static noise. This suggests a central mechanism that requires coherent modulation to extract kinetic edges.

Four Lines of Evidence that Central Mechanisms Mediate Stabilized Image Phenomena

(1) Effects of Binocular and Other Sensory Interactions

(a) Stabilized monocular image fading can be reversed by similarly patterned stimulation of the other eye (Cohen, 1961), especially if placed on corresponding points of the stabilized and unstabilized retinae (effectiveness drops off monotonically with retinal disparity). In dichoptic presentation, modulating the background eye slows fading of the image in the stabilized eye (Gerling & Spillmann, 1987). (b) Conversely, fading of a stabilized image makes unstabilized stimuli in the other eye less visible (Krauskopf & Riggs, 1959). (c) Even after a stabilized image has disappeared in one eye, it can be combined with an unstabilized duplicate moving image in the other eye, to yield a sensation of motion in depth (Crane, 1994). (d) Some stabilized images induce an abrupt absolute blindness (Billock, Gleason, & Tsou, 2001; Ditchburn, 1973). Both blindness and recovery are binocularly simultaneous, implying a central mechanism. (e) Visibility of stabilized images can be maintained by listening to auditory stimuli; the effect decays to baseline in about 15 min if the stimulation is not varied, suggesting a role for attention (Ditchburn, 1973).

(2) Fragmentation

Davies (1973), Evans (1965), and Pritchard, Heron, and Hebb (1960) find complex stabilized images dis/reappear, not as wholes but in fragmented forms (for illustrations of fragmentation percepts, see Billock & Tsou, 2004). This image fragmentation is not explained by the properties of peripheral neurons (e.g., fragmentation can eliminate contours that are present only in a binocular image, but not in the separate retinal images; Evans & Wells, 1967). Moreover, image fragments that wax and wane together obey Gestalt-like rules (Evans, 1965, 1967): (a) Short lines appear and disappear as a unit. Fragmentation probability increases with line length; for foveal vision fragmentation is likely for line lengths of roughly 45 arc min (an order of magnitude larger than foveal retinogeniculate receptive fields), suggesting a range for the cortical mechanisms involved. (b) Disappearance and reappearance of parallel lines are correlated. (c) Random patterns are more unstable than meaningful ones. (d) Angular patterns are more fragmented and perceptually unstable than rounded patterns. (e) Fragmentation and fading in one region of a field is strongly affected by activity in neighboring regions. (f) Complex patterns are more likely to fragment than simple ones. Simulations show that Gestalt-like grouping rules are emergent properties of cooperative segmentation and binding networks (Horn & Opher, 2000; Wang & Terman, 1997; Sporns, Tononi, & Edelman, 1991). Interestingly, some models of synchronized segmentation tend to fragment images if the spatial properties of the model do not pool over a large enough set of units to discount noisy stimulus inputs (Terman & Wang, 1995). Similar fragmentation is reported by some subjects with migraine (mosaic vision; Sacks, 1995) and amblyopia (especially for high spatial frequency stimuli; Hess, Field, & Watt, 1990); a link between amblyopia and neural synchronization is suggested by Roelfsema, König, Engel, Sireteanu, and Singer (1994) finding that cortical cells are normal in spatial frequency selectivity, but cells driven by the amblyopic eye do not synchronize well, especially for high spatial frequency stimuli.

(3) Filling-in of Stabilized Images

If a broken figure is stabilized and fades—then regenerates—the reappearance of the figure is often marked by completion of the break (Ditchburn, 1973). Moreover, Cardu, Gilbert, and Stabel (1971) find that objects made up of dashed lines often (45% of the time) exhibit completion just before fading. This effect resembles the completion of images broken by scotomas and is compatible with binding algorithms that incorporate cooperative interactions.

(4) Stabilization Effects on EEG Correlated to Perceptual Phenomena

Although no stabilization experiment has looked for synchronization of γ-band activity, several studies find that the relative level of α-rhythm activity increases when stabilized images disappear or when perceptual blanks occur during Ganzfeld viewing (Keesey & Nichols, 1967, 1969; Lehmann, Beeler, & Fender, 1967, Lehmann, Beeler, & Fender, 1965; Evans & Smith, 1964). Following stabilization, power in the EEG α-band

(9–10 Hz) rises about 0.7–1.0 sec before target disappearance and higher frequencies are suppressed until about 1 sec prior to reappearance of target structures (Keesey & Nichols, 1967; Lehmann et al., 1965).

Overview of Stabilized Image Studies: Only a Central Explanation Will Do

The textbook account of stabilized image perception is that fading is due to the transient response properties of some retinogeniculate units. That analysis cannot account for counterintuitively rapid fading of stimuli that favor highly sustained parvocellular mechanisms. Nor can it account for any of the explicitly cortical effects discussed above. Of these, image fading and fragmentation are consistent with known failure modes of segmentation networks that require a balance of competitive and cooperative interactions; fading can stem from relative weakness of competitive interactions; fragmentation can stem from weakness of cooperative interactions. The Gestalt properties shown in stabilized image studies correspond to emergent properties of cooperative/competitive segmentation networks (Horn & Opher, 2000; Wang & Terman, 1997; Sporns et al., 1991). As early as 1960, Pritchard et al. argued that fragmentation and partial reappearance of faded stabilized images is evidence for the dynamic formation of Hebbian cell assemblies—essentially the same argument advanced in modern binding theory. Here, we argue that binding of units responding to nonstabilized images is facilitated by the common information in each form and motion mechanism induced by eye movements perturbing the retinal image, allowing motion and form-based segmentation networks to reinforce each other.

DISCUSSION

Summary of Reasoning

At this point, it is useful to summarize our reasoning and the evidence that supports it so that we can lay out its strengths and weaknesses, and point out where future work needs to be done:

1. Feng (2000) and Horn et al. (1991) showed that a preexisting correlation between mechanisms facilitates synchronizing those mechanisms' activity, a form of feature binding. In Horn et al.'s and Feng's work, these initial correlations stem from unspecified noise. What we add here is an explicit physiological source for such signals (crosstalk induced by cortical demultiplexing of LGN signals), which results in stimulus-driven correlations between cortical mechanisms for achromatic form and color, and between achromatic form and motion.

2. Several investigators point out that feature-binding between mechanisms responding to common parts of an image can improve part-binding/segmentation (Schillen & König, 1994; Finkel & Edelman, 1989; Poggio et al., 1988), a process called crossmodal reinforcement.

3. Other investigators point out that some part-binding (segmentation) models can fail in interesting ways, for example, fragmentation of images (Terman & Wang, 1995; Marr et al., 1978). What we add to this is a discussion of various ways a segmentation network could fail (depending on whether cooperative or competitive interactions were inadequate).

4. We point out that some odd percepts (e.g., fragmentation of stabilized images and melting of equiluminant images) closely resemble these expected failure modes of segmentation models. We suggest that these part-binding failures may be due to loss of crossmodal reinforcement (Point 2) that under normal circumstances is available and relatively easy to bind because of stimulus-driven correlations (Point 1).

Modeling—What Remains To Be Done

Although each point in the above analysis is supported by experiments or mathematical models, it would be desirable to supplement this analysis with a formal integrated simulation and with a set of experimental predictions to be tested. We provide some such predictions in the sections below. A formal integrated simulation lies beyond the scope of this article, but we can lay out what would be required to implement one. For simplicity, we discuss form and color binding here—the steps required for form and motion binding are quite similar. (a) Start with a segmentation model, like Terman and Wang's (1995) that fails in one of the interesting ways described in Point 4, and manipulate the balance of competition and cooperation in the model until it so fails. Two copies of this segmentation network are created, although they need not be symmetrical (only failure in the color segmentation system is being studied in this example). (b) Create two sets of input mechanisms to drive the networks in (a)—an array of color units obeying Equation 9 and an array of achromatic form units obeying Equation 8. To parameterize these equations, the S_e, T_e, S_i, and T_i functions describing the spatial and temporal excitatory and inhibitory responses of P cells need to be specified using psychophysical or physiological data. Suitable functions (derived psychophysically) can be found in Kelly (1989) and Burbeck and Kelly (1980). Under ordinary (non-equiluminant) conditions, there should be a correlation between the color inputs described by Equation 9 and the achromatic units described by Equation 8. This correlation is implicit in the crosstalk terms of these equations, but can be made explicit using Equation 10. Now run the segmentation networks so that they synchronize for normal stimuli (stimuli that contain both hue and luminance information). (c) Set the stimuli to be equiluminant. The response of the achromatic form-

driven network should become disorganized (because it is being driven by noise), and the chromatic network should desynchronize from it. Segmentation mistakes in the unreinforced chromatic network should now occur.

Combining Equiluminance and Stabilization—Effects and Predictions

Many of the binding failures discussed above seem connected. Equiluminant border melting has obvious similarities to border fading in image stabilization. Similarly, the ''jazziness'' of equiluminant borders may be a fast time-scale analog to the fragmentation and dis/reappearance of some stabilized images. The completion of gaps in reappearances of stabilized images resembles filling-in phenomena seen for some equiluminous images. In general, nulling the signal to a reinforcing channel should be detrimental to the channel it reinforced: The reinforcement is lost and lost information is replaced with mere noise. Significantly, for subjects who have one of the color, motion, and achromatic form systems disrupted, interrupting another system results in a more dramatic effect. For example, Sacks and Wasserman (1987) find that for an achromatopsic observer, boundaries between areas dissolved during periods of steady fixation (see below for analogous findings in color normals). Similarly, stimuli that stimulate only the b–y pathway (tritanopic stimuli) are particularly susceptible to dissolution of contours (Cavanagh, 1991). If the visual system requires two cooperating systems to achieve segmentation, then we would predict that combining equiluminance and stabilization should be especially devastating. In fact, steadily fixated colored stimuli do fade more readily when equiluminant to their surroundings (Livingstone & Hubel, 1987; Buck et al., 1977; Krauskopf, 1967). Still more dramatic effects are obtained by combining retinal stabilization and equiluminance. Billock et al. (2001) and Crane and Piantanida (1983) report that if a stabilized red/green or blue/yellow bipartite field is viewed through an unstabilized aperture, the colors flow and mix across the faded border. Some observers perceive uniform reddish greens or yellowish blues, in violation of color opponency. Other observers report formation of fine color textures or unstable islands of one color continuously forming and dissolving in a sea of the other color. Billock et al. found this spurious segmentation occurs only when there is a luminance difference between the two stabilized colored fields. Moreover, while melting together of adjacent equiluminant fields is normally seen only for tritanopic pairs, Billock et al. found that the phenomenon occurs for any stabilized color pair. (Related and extreme cases of binding failures are found in some neurological disorders; Friedman-Hill, Robertson, & Treisman, 1995; Sacks, 1995; Zeki, 1993; Critchley, 1965).

Given our reasoning and these preliminary results, a test of our reinforcement hypothesis would be to measure both perceptual segmentation and electrophysiological measures of binding while equiluminance and stabilization are manipulated. The psychophysics is straightforward. For experiments on equiluminance, an image is made of at least two colors, and radiance ratios of the colors are varied through the range in which the equiluminance ratio must lie. Whether loss of neural binding signatures occurred at equiluminance can be determined by measuring figure and background luminosities using standard psychophysical techniques (Cavanagh, 1991; Wyszecki & Stiles, 1982). Specific information should be gathered with respect to perceptual state (e.g., border instability, border melting, surface delinking). Similarly, image stabilization is straightforward, albeit technically challenging and expensive (Crane, 1994); useful results can often be had with afterimages or steady fixation (Ditchburn, 1973). Electrophysiological states should be correlated with the specific perceptual state elicited (e.g., fragmentation, fading, etc.). If a stabilization system is available, the degree of stabilization can be manipulated by moving the image independently of eye movements or by manipulating the gain of the eye-movement $\leftrightarrow$ image-movement feedback loop (Crane, 1994). Finding what neural signatures to use (and how to test for them noninvasively in responding humans) is more challenging. Since γ-band activity is implicated in binding, the loss of higher-frequency EEG power during perceptual losses in image stabilization is suggestive (Keesey & Nichols, 1967; Lehmann et al., 1965). There are two limitations with this approach: Some sources of γ-band power may be unrelated to binding and some binding failures may represent incorrect bindings and therefore not be identifiable by EEG. Recent experiments offer help on both counts, by studying gestalt perception (pop-out) of fragmented images hidden in camouflage. This fragmentation-to-order shift is complementary to the order-to-fragmentation effects we wish to study. Subjects had γ-band activity whether they saw a coherent image pop-out or not, but for coherent percepts the outputs of widely spaced electrodes became more correlated (as would be expected for synchronization; Gruber et al., 2002) and a different kind of γ-band activity ''the induced gamma response'' was recorded. On every physically identical trial, there was a γ-band response phase-locked to the stimulus, but only on trials where the gestalt percept was obtained was there another EEG signature, consisting of γ-band bursts in variable phase to the stimulus (Gruber et al., 2002; Tallon-Baudry & Bertrand, 1999; Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996). We therefore predict that EEG signatures like ''induced gamma'' will be reduced during binding failures induced by retinal stabilization (especially fragmentation) and equiluminance (e.g., loss of Gestalt symmetry in Glass patterns; Cavanagh, 1991). Moreover, we posit that these signatures should be further reduced by combining equilumi-

nance with stabilization, and should be monitored when color boundaries collapse and forbidden colors are perceived (Billock et al., 2001). An easier variation of this experiment could be done with steady fixation on a minimized border (Buck et al., 1977); under these conditions, some color boundaries collapse, creating uniform color mixtures.

Acknowledgments

We thank Randolph Blake, Angela Brown, Patrick Cavanagh, Viktor Jirsa, Daniel Levine, Lynn Olzak, Wolf Singer, DeLiang Wang, and Scott Watamaniuk for critical readings of the manuscript. Special thanks to J. A. Scott Kelso for suggesting this problem.

Reprint requests should be sent to V. A. Billock, General Dynamics, Suite 200, 5200 Springfield Pike, Dayton, OH 45431, USA, or via e-mail: Vince.Billock@wpafb.af.mil.

REFERENCES

Benardete, E. A., Kaplan, E., & Knight, B. W. (1992). Contrast gain control in the primate retina: P cells are not X-like, some M-cells are. *Visual Neuroscience, 8,* 483–486.

Billock, V. A. (1996). Consequences of retinal color coding for cortical color decoding. *Science, 274,* 2118–2119.

Billock, V. A. (1995). Cortical simple cells can extract achromatic information from the multiplexed chromatic and achromatic signals in the parvocellular pathway. *Vision Research, 35,* 2359–2369.

Billock, V. A. (1991). The relationship between simple and double opponent cells. *Vision Research, 31,* 33–42.

Billock, V. A., Gleason, G. A., & Tsou, B. H. (2001). Perception of forbidden colors in retinally stabilized equiluminant images: An indication of softwired cortical color opponency? *Journal of the Optical Society of America A, 18,* 2398–2403.

Billock, V. A., & Tsou, B. H. (2004). What do catastrophic visual binding failures look like? *Trends in Neurosciences, 27,* 84–89.

Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review, 83,* 1–36.

Buck, S. L., Frome, F., & Boynton, R. M. (1977). Initial distinctness and subsequent fading of minimally distinct borders. *Journal of the Optical Society of America, 67,* 1126–1128.

Burbeck, C. A., & Kelly, D. H. (1980). Spatiotemporal characteristics of visual mechanisms: Excitatory–inhibitory model. *Journal of the Optical Society of America, 70,* 1121–1126.

Burt, P. J. (1987). The interdependence of temporal and spatial information in early vision. In M. A. Arbib (Ed.), *Vision, brain, and cooperative computation* (pp. 263–277). Cambridge: MIT Press.

Campbell, F. W., Hopwell, E. R., & Johnstone, J. R. (1978). A comparison of threshold and superthreshold appearance of gratings with components in the low and high spatial frequency range. *Journal of Physiology, 284,* 193–201.

Cardu, B., Gilbert, M., & Stabel, M. (1971). The influence of peripheral and central factors on the way that stabilized images disappeared. *Vision Research, 11,* 1337–1343.

Cavanagh, P. (1991). Vision at equiluminance. In J. J. Kulikowski, V. Walsh & I. J. Murray (Eds.), *Limits of vision* (pp. 234–250). Boca Raton: CRC Press.

Cohen, H. B. (1961). The effect of contralateral visual stimulation on visibility with stabilized retinal images. *Canadian Journal of Psychology, 15,* 212–219.

Coppola, D., & Purves, D. (1996). The extraordinarily rapid disappearance of entopic images. *Proceedings of the National Academy of Sciences, U.S.A., 93,* 8001–8004.

Crane, H. D. (1994). The Purkinje image eyetracker, image stabilization, and related forms of stimulus manipulation. In D. H. Kelly (Ed.), *Visual science and engineering* (pp. 15–89). New York: Marcel Dekker.

Crane, H. D., & Piantanida, T. P. (1983). On seeing reddish green and yellowish blue. *Science, 221,* 1078–1080.

Critchley, M. (1965). Acquired anomalies of colour perception of central origin. *Brain, 88,* 711–724.

Davies, P. (1973). The role of central processes in the perception of visual after-images. *British Journal of Psychology, 64,* 325–338.

de Monasterio, F. M., & Gouras, P. (1975). Functional properties of ganglion cells of the rhesus monkey retina. *Journal of Physiology, 251,* 167–195.

De Valois, R. L., & De Valois, K. K. (1988). *Spatial Vision.* New York: Oxford University Press.

De Valois, R. L., & Pease, P. L. (1971). Contours and contrast: Responses of monkey lateral geniculate cells to luminance and color figures. *Science, 171,* 694–696.

Ditchburn, R. W. (1973). *Eye movements and visual perception.* Oxford: Clarendon.

Evans, C. R. (1967). Further studies of pattern perception and a stabilized retinal image. *British Journal of Psychology, 58,* 315–327.

Evans, C. R. (1965). Some studies of pattern perception using a stabilized retinal image. *British Journal of Psychology, 56,* 121–133.

Evans, C. R., & Smith, G. K. (1964). Alpha-frequency of electroencephalogram and a stabilized retinal image. *Nature, 204,* 303–304.

Evans, C. R., & Wells, A. M. (1967). Fragmentation phenomena associated with binocular stabilization. *British Journal of Physiological Optics, 24,* 45–50.

Feng, J. (2000). Synchronization driven by correlated inputs. *Neurocomputing, 32,* 371–387.

Finkel, L. H., & Edelman, G. M. (1989). Integration of distributed cortical systems by reentry: A computer simulation of interactive functionally segregated visual areas. *Journal of Neuroscience, 9,* 3188–3208.

Friedman-Hill, S. R., Robertson, L. C., & Treisman, A. (1995). Parietal contributions to visual feature binding: Evidence from a patient with bilateral lesions. *Science, 269,* 853–855.

Gerling, J., & Spillmann, L. (1987). Duration of visual afterimages on modulated backgrounds: Postreceptoral processes. *Vision Research, 27,* 521–527.

Gray, C. M. (1999). The temporal correlation hypothesis of visual feature integration: Still alive and well. *Neuron, 24,* 31–47.

Gregory, R. L. (1977). Vision with isoluminant colour contrast. *Perception, 6,* 113–119.

Grossberg, S., & Wyse, L. (1992). Figure–ground separation of connected scenic figures: Boundaries, filling-in and opponent processing. In G. A. Carpenter & S. Grossberg (Eds.), *Neural networks for vision and image processing* (pp. 161–194). Cambridge. MA: MIT Press.

Gruber, T., Müller, M. M. & Keil, A. (2002). Modulation of induced gamma band responses in a perceptual learning task in the human EEG. *Journal of Cognitive Neuroscience, 14,* 732–744.

Hess, R. F., Field, D. J., & Watt, R. J. (1990). The puzzle of amblyopia. In C. Blakemore (Ed.), *Vision: coding and efficiency* (pp. 267–280). Cambridge: Cambridge University Press.

Horn, D., & Opher, I. (2000). Temporal segmentation and binding in oscillatory neural systems. In D. S. Levine, V. R. Brown, & V. T. Shirey (Eds.), *Oscillations in neural systems* (pp. 201–216) Mahwah, NJ: Lawrence Erlbaum.

Horn, D., Sagi, D., & Usher, M. (1991). Segmentation, binding and illusory conjunctions. *Neural Computation, 3,* 510–525.

Humanski, R. A., & Wilson, H. R. (1993). Spatial frequency adaptation: Evidence for a multiple channel model of short-wavelength-cone spatial vision. *Vision Research, 33,* 665–675.

Ingling, C. R., Jr., & Grigsby, S. S. (1990). Perceptual correlates of magnocellular and parvocellular channels: Seeing form and depth in afterimages. *Vision Research, 30,* 823–828.

Ingling, C. R., Jr., & Martinez, E. (1983). The relationship between spectral sensitivity and spatial sensitivity for the primate r–g X-cell channel. *Vision Research, 23,* 1495–1500.

Ingling, C. R., Jr., & Martinez-Uriegas, E. (1985). The spatiotemporal properties of the r–g X-cell channel. *Vision Research, 25,* 33–38.

Keesey, U. T., & Nichols, D. J. (1969). Changes induced in stabilized image visibility by experimental alteration of the ongoing EEG. *Electroencephalography and Clinical Neurophysiology, 27,* 248–257.

Keesey, U. T., & Nichols, D. J. (1967). Fluctuations in target visibility as related to the occurrence of the alpha component of the electroencephalogram. *Vision Research, 7,* 859–879.

Kelly, D. H. (1989). Opponent-color receptive-field profiles determined from large-area psychophysical measurements. *Journal of the Optical Society of America A, 6,* 1784–1793.

Kelly, D. H. (1983). Spatiotemporal variation of chromatic and achromatic contrast thresholds. *Journal of the Optical Society of America, 73,* 742–750.

Kelly, D. H. (1981). Disappearance of stabilized chromatic gratings. *Science, 214,* 1257–1258.

Kelly, D. H., & Martinez-Uriegas, E. (1993). Measurements of chromatic and achromatic afterimages. *Journal of the Optical Society of America A, 10,* 29–37.

Kingdom, F. A. A., & Mullen, K. (1995). Separating colour and luminance information in the visual system. *Spatial Vision, 9,* 191–219.

King-Smith, P. E., Rosten, J. G., & Alverez, S. L. (1980). Human vision without tonic ganglion cells? In G. Verriest (Ed.), *Colour Vision Deficiencies V* (pp. 99–105). Bristol: Hilger.

Kotulak, J. C., Morse, S. E., & Billock, V. A. (1995). Red–green opponent channel mediation of control of human ocular accommodation. *Journal of Physiology, 482,* 697–703.

Krauskopf, J. (1967). Heterochromatic stabilized images. *American Journal of Psychology, 80,* 634–637.

Krauskopf, J., & Riggs, L. A. (1959). Interocular transfer in the disappearance of stabilized images. *American Journal of Psychology, 72,* 248–252.

Lehmann, D., Beeler, G. W., & Fender, D. H. (1967). EEG responses to light flashes during the observation of stabilized and normal retinal images. *Electroencephalography and Clinical Neurophysiology, 22,* 136–142.

Lehmann, D., Beeler, G. W., & Fender, D. H. (1965). Changes in patterns of the human electroencephalogram during fluctuations of perception of stabilized retinal images. *Electroencephalography and Clinical Neurophysiology, 19,* 336–343.

Levine, D. S. (2000). *Introduction to neural and cognitive modeling.* Hillsdale, NJ: Erlbaum.

Liebmann, S. (1927). Uber das Verhalten farbiger Formen bei Helligkeitsgleichheit von Figur und Grund. *Psychologische Forschung, 9,* 300–353.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for perception of form, color, movement and depth. *Journal of Neuroscience, 7,* 3416–3468.

Livingstone, M. S., & Hubel, D. H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience, 4,* 309–356.

Lu, C., & Fender, D. H. (1972). The interaction of color and luminance in stereoscopic vision. *Investigative Ophthalmology and Visual Science, 11,* 482–490.

Marr, D., Palm, G., & Poggio, J. (1978). Analysis of a cooperative stereo algorithm. *Biological Cybernetics, 28,* 223–239.

Merigan, W. H., & Maunsell, J. H. R. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience, 16,* 369–402.

Mullen, K. T., & Kingdom, F. A. A. (1991). Colour contrast in form perception. In P. Gouras (Ed.), *The perception of colour* (pp. 198–217). Boca Raton: CRC Press.

Poggio, T., Gamble, E. B., & Little, J. J. (1988). Parallel integration in visual modules. *Science, 242,* 436–440.

Pritchard, R. M., Heron W., & Hebb, D. O. (1960). Visual perception approached by the method of stabilized images. *Canadian Journal of Psychology, 14,* 67–77.

Regan, D. (2000). *Human perception of objects: Early visual processing of spatial form defined by luminance, color, texture, motion and binocular disparity.* Sunderland, MA: Sinauer.

Roelfsema, P. R., Engel, A. K., König, P., & Singer, W. (1996). The role of neuronal synchronization in response selection: A biologically plausible theory of structured representations in the visual cortex. *Journal of Cognitive Neuroscience, 8,* 603–625.

Roelfsema, P. R., König, P., Engle, A. K., Sireteanu, R., & Singer, W. (1994). Reduced synchronization in the visual cortex of cats with strabismic amblyopia. *European Journal of Neuroscience, 6,* 1645–1655.

Sacks, O. (1995). *Migraine.* London: Picador.

Sacks, O., & Wasserman, R. (1987). The painter who became color blind. *New York Review of Books, 34,* 25–33.

Schillen, T. B., & König, P. (1994). Binding by temporal structure in multiple feature domains of an oscillatory neuronal network. *Biological Cybernetics, 70,* 397–405.

Schiller, P., Logothesis, N. K., & Charles, E. R. (1990). Functions of the colour-opponent and broad-band channels of the visual system. *Nature, 343,* 68–70.

Shadlen, M. N., & Movshon, J. A. (1999). Synchrony unbound: A critical evaluation of the temporal binding hypothesis. *Neuron, 24,* 67–77.

Singer, W. (1999). Neuronal synchrony: A versatile code for the defination of relations? *Neuron, 24,* 49–65.

Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience, 18,* 555–586.

Spillmann, L., & Kurtenbach, A. (1992). Dynamic noise backgrounds facilitate target fading. *Vision Research, 32,* 1941–1946.

Sporns, O., Tononi, G., & Edelman, G. M. (1991). Modeling perceptual grouping and figure–ground segregation by means of active reentrant connections. *Proceedings of the National Academy of Sciences, U.S.A., 88,* 129–133.

Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity and its role in object representation. *Trends in Cognitive Sciences, 3,* 151–162.
Tallon-Baudry, C., Bertrand, O., Delpuech, C. & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 hz visual responses in human. *Journal of Neuroscience, 16,* 4240–4249.
Terman, D., & Wang, D. L. (1995). Global competition and local cooperation in a network of neural oscillators. *Physica D, 81,* 148–176.
Thorell, L. G., De Valois R. L., & Albrecht, D. G. (1984). Spatial mapping of monkey V1 cells with pure color and luminance stimuli. *Vision Research, 24,* 751–769.
Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology, 6,* 171–178.
von der Malsburg, C. (1999). The what and why of binding: The modeler's perspective. *Neuron, 24,* 95–104.
von der Malsburg, C. (1995). Binding in models of perception and brain function. *Current Opinion in Neurobiology, 5,* 520–526.
von der Malsburg, C. (1981). The correlation theory of brain function. *Max Planck Institute for Biophysical Chemistry Report* 81–82.
von der Malsburg, C., & Buhmann, J. (1992). Sensory segmentation with coupled neural oscillators. *Biological Cybernetics, 67,* 233–242.
Wang, D. L., & Terman, D. (1997). Image segmentation based on oscillatory correlation. *Neural Computation, 8,* 805–836.
Wiesel, T. N., & Hubel, D. (1966). Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *Journal of Neurophysiology, 29,* 1115–1156.
Wyszecki, G., & Stiles, W. S. (1982). *Color science.* New York: Wiley.
Zeki, S. (1993). *A Vision of the Brain.* Oxford: Blackwell.