Perception & Psychophysics
2000, 62 (6), 1160-1181

Lightness from contrast: A selective integration model

WILLIAM D. ROSS
MIT Lincoln Laboratory, Lexington, Massachusetts

and

LUIZ PESSOA
Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

As has been observed by Wallach (1948), perceived lightness is proportional to the ratio between the luminances of adjacent regions in simple disk–annulus or bipartite scenes. This psychophysical finding resonates with neurophysiological evidence that retinal mechanisms of receptor adaptation and lateral inhibition transform the incoming illuminance array into local measures of luminance contrast. In many scenic configurations, however, the perceived lightness of a region is not proportional to its ratio with immediately adjacent regions. In a particularly striking example of this phenomenon, called White's illusion, the relationship between the perceived lightnesses of two gray regions is the opposite of what is predicted by local edge ratios or contrasts. This paper offers a new treatment of how local measures of luminance contrast can be selectively integrated to simulate lightness percepts in a wide range of image configurations. Our approach builds on a tradition of edge integration models (Horn, 1974; Land & McCann, 1971) and contrast/filling-in models (Cohen & Grossberg, 1984; Gerrits & Vendrik 1970; Grossberg & Mingolla, 1985a, 1985b). Our selective integration model (SIM) extends the explanatory power of previous models, allowing simulation of a number of phenomena, including White's effect, the Benary Cross, and shading and transparency effects reported by Adelson (1993), as well as aspects of motion, depth, haploscopic, and Gelb induced contrast effects. We also include an independently derived variant of a recent depthful version of White's illusion, showing that our model can inspire new stimuli.

In everyday experience, surface color constancy is an effortless achievement of the visual system. That is, despite variations in lighting and movement or displacement of objects across visual contexts, object color appears to a large extent to remain constant. For example, consider the appearance of a teapot in a familiar kitchen scene. Large daily variations in the illumination of the kitchen, including shadowing, do not alter the apparent surface color of the teapot (Type I, or illumination-independent constancy). What is more, the teapot remains the same apparent color, despite being placed at different locations, whether on a red table cloth or on a white counter (Type II, or background-independent constancy). Color constancy refers, then, to the fact that surface color remains largely constant, despite changes in the intensity and composition of the light reflected to the eyes from both the object itself and from surrounding objects.[1] This paper is specifically concerned with the achromatic, or black-to-white, dimension of perceived surface color (perceived reflectance), which is often referred to as *lightness*. As do others, we assume that *brightness*—the perceived luminance of a region of space—can be indistinguishable from and thus, at times, equivalent to lightness for simple flat displays (see Fiorentini, Baumgartner, Magnussen, Schiller, & Thomas, 1990)

L.P. was funded by CNPq/Brazil Grant 520419/96-0. W.D.R. was supported in part by the Whitaker Foundation. The authors thank Alan Gilchrist, Robert O'Shea, and the other reviewers for valuable discussions concerning the presentation of the model. Correspondence concerning this article should be addressed to W. D. Ross, Machine Intelligence Group, MIT Lincoln Laboratory, Lexington, MA (e-mail: bross@ll.mit.edu).

The Ratio Principle

Wallach (1948) conducted a seminal lightness constancy experiment. He showed that for simple disk– annulus stimuli, perceived lightness is proportional to the ratio of disk luminance to annulus luminance. Wallach's results offered an explanation of lightness constancy under changes in uniform scene illumination (Type I constancy)—sometimes referred to as Wallach's *ratio principle*. According to this principle, lightness percepts correspond to measures of the ratio between local luminance and nearby luminances, rather than to direct measures of local luminance. The ratio principle may be the means by which the visual system unconfounds the dual effects of illumination and surface reflectance on local luminance. Ratios between nearby luminances tend to remain constant as overall illumination levels vary, so this strategy achieves a certain degree of illumination-independent lightness constancy (Arend & Goldstein, 1987; Jacobsen & Gilchrist, 1988).

Later perceptual studies further supported the correspondence between relative luminance and lightness mea-

Copyright 2000 Psychonomic Society, Inc.

surement in early visual processing (see Gilchrist, 1994). Whittle and Challands (1969) had subjects perform brightness matches in a haploscopic display paradigm. In a haploscopic display, each eye sees a separate image; in each of these, a patch (the patches are identical) appears on a different background. When the stimulus is binocularly fused, the two patches appear to be on the same background but have different lightnesses (as in a standard simultaneous contrast display). The striking result was that when the test-patch luminance consisted of a decrement relative to its background, subjects always matched it with relative decrements. Likewise, increments were always matched with increments, but never increments with decrements. These results support the notion that lightness is defined in relational terms.

Neurophysiological studies confirm that cells at early stages of the visual system do seem to encode local luminance contrast (Shapley & Enroth-Cugell, 1984) through the mechanisms of light adaptation and lateral inhibition (Walraven, Enroth-Cugell, Hood, MacLeod, & Schnapf, 1990). Throughout this paper, we employ the phrase *luminance contrast*, or just *contrast*, when referring to a physical property of a display. When the word is employed in a different sense, we employ modifiers, such as *simultaneous contrast*—which is a perceptual phenomenon, or *contrast measures*—which refer to the results of model or neural computations. An image showing luminance contrast measures, computed through modeling the center–surround antagonistic retinal mechanisms, is shown in Figure 1B: Middle gray denotes scenic regions without associated contrast information; brighter and darker shades code positive and negative luminance transitions, respectively, for the stimulus shown in Figure 1A.

A number of past computational investigations have taken seriously the implication that early vision encodes luminance contrasts, and some have achieved partial success in modeling how contrast measures might be used by the brain to generate lightness percepts. Filling-in models (e.g., Gerrits & Vendrik, 1970; Grossberg & Todorović, 1988) show that the averaging of local contrast ratios within bounded regions can correctly determine perceived brightness for a number of experimental scenes. Integration models (e.g., Horn, 1974; Land & McCann, 1971) measure local contrasts at an initial differentiation stage whose aim is to recover local differences or ratios; after thresholding, these can be integrated across a scene to accurately model lightness percepts in a number of scenes.

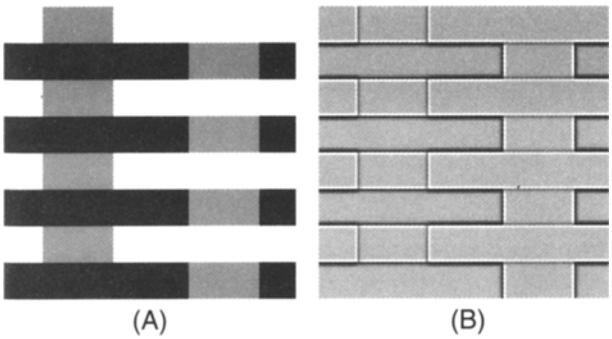

Figure 1. (A) White's illusion. The patches are colored the same gray level. The patches within the black stripes typically appear lighter than the patches within the white stripes. (B) Retinal contrast measures computed through adaptation and lateral inhibition only carry information about luminance gradients (simulated *on–off* channel responses are shown, with middle gray coding no *on* or *off* cell activation, and brighter and darker shades coding activities in the *on* and *off* channels, respectively).

Psychophysical, neurophysiological, and computational studies all support the plausibility of some version of Wallach's (1948) ratio principle as a significant insight into early visual processing and lightness perception in simple scenes. In general, however, local ratios prove insufficient to account for lightness perception. Striking contradictions of the ratio principle have been reported in more complex structured scenes. These contradictions are apparent in studies of the role of 3-D spatial layout and illumination arrangement on lightness perception (Gilchrist, 1977, 1980; Hochberg & Beck, 1954). In 1977, they led Gilchrist to propose that the data are better characterized by a *coplanar ratio hypothesis* stating that ratios between regions within the same depth plane are more significant in determining lightness than are ratios between depth planes. However, contradictions of the ratio principle have also been reported in studies of flat stimulus configurations (e.g., Benary's [1924] cross, and White's [1979] illusion, which is shown in Figure 1A) and edge-integration effects (Arend, Buehler, & Lockhead, 1971; Reid & Shapley, 1988; Shapley & Reid, 1985), as well as several display configurations involving changes in background (or surround) luminance (Arend & Spehar, 1993; Gilchrist, 1988; Land & McCann, 1971; Whittle, 1992; Whittle & Challands, 1969).

Selective Integration

A remarkable example of a violation of the ratio principle is White's effect. Although the gray patches in Figure 1A have the same physical reflectance, the ones on the left are typically perceived to be darker than the ones on the right. White's effect is considered a puzzling violation of the predictions of local luminance ratios since the contour length of the gray patches is larger for the noncoaxial stripes. Local contrast measures (computed through processes sensitive to luminance gradients, such as in Figure 1B) would predict that the gray patches coaxial with the black stripes would appear darker than the ones coaxial with the white stripes. This example highlights the *context-sensitive* transformation from contrast measures to lightness that is accomplished by the visual system.

It has long been recognized that to determine lightness accurately from local luminance ratios in complex scenes, edge contrasts must be able to influence not only the determination of lightness within adjacent bounded regions, but also the determination of lightness in nonadjacent regions (e.g., Arend et al., 1971). Local ratios are by definition measures of local luminance relationships. Only when taken as a whole set do they embody information about relationships between widely separated regions. Integration models of lightness perception have sought to formalize how the determination of scenic lightness can be made on the basis of the whole set of contrast measures in a scene (Arend & Goldstein, 1987; Blake, 1985; Horn, 1974; Land & McCann, 1971). An important contribution from this area has been the discovery that all contrast measures are not weighted equally in the determination of lightness. For example, Gilchrist and colleagues have suggested that lightness constancy depends on discounting the ratios at illumination edges, such as shadow borders, during the integration process (Gilchrist, Delman, & Jacobsen, 1983).

In the rest of this paper, we present a model that builds on previous edge integration and filling-in models. Our selective integration model (SIM) treats contrast measures as *constraints* on local lightness relationships, as in other integration models. Lightnesses are determined in such a way that the lightness ratio between any two regions tends to be consistent with an integration of all the edge contrasts separating the two regions. Integration within SIM, however, is *selective*. That is, the weighting of edge contrasts in determining the lightness of a region is sensitive to visual groupings or contexts. In particular, those contrasts within a context are emphasized in the determination of lightness percepts within that context, whereas contrasts between contexts are partially reduced. We show that SIM accounts for a number of phenomena that have proven difficult for previous models to explain, including White's effect, the Benary cross, and the shading and transparency effects reported by Adelson (1993), as well as certain aspects of motion, depth, haploscopic, and Gelb contrast effects.

THE SELECTIVE INTEGRATION HYPOTHESIS: WHITE'S EFFECT

All edge contrasts do not signal the same type of scenic events. Some outline regions of uniform coloration on surfaces; others define the borders between objects, surfaces, or illumination contexts. Selectively reducing integration across these borders helps prevent the mixing of information across contexts. Selective integration can thus act to reduce the impact of illumination context differences, such as those introduced by shadows, spotlights, shading, and transparency, as well as depth changes, on perceived lightness. Furthermore, selective integration acts to emphasize the stable luminance relationships within objects over those variable relationships between an object and changing backgrounds (background-independent constancy).

The proposal that the visual system specifies lightness through context-sensitive processes of integration is, of course, not new. To mention one example, Gilchrist (1977) showed that in 3-D scenes, the lightness of a given region is determined predominantly in relation to other coplanar regions, and not by equally weighted relations to all retinally adjacent regions. We suggest that other contradictions of the ratio principle can also be explained by context-selective integration mechanisms.

White's Effect

In order to explore this possibility, consider White's effect. Perhaps the simplest segmentation of the stimulus shown in Figure 1 would be to separate the gray patches from the horizontal stripes immediately above and below them, which may appear to occlude them (in this version of the stimulus, this segmentation is most obvious for the black bars). This segmentation divides the stimulus into horizontal stripes. If the lightness of the gray patches is given predominantly by their relation to their neighbors in a horizontal direction, the patches coaxial with white stripes should appear darker than the patches coaxial with black stripes, as is perceived.

We tested the above segmentation hypothesis of White's illusion by manipulating scenic segmentation without significantly altering the defining luminance relationships in the image projections. This was accomplished by manipulating depth in such a way that both consistent and inconsistent depth groupings were produced, where by *consistent*, we mean a grouping similar to that assumed in the analysis of the flat stimulus of Figure 1 and other groupings are, thus, *inconsistent*. Figure 2 shows stereo images that induce a consistent depth grouping.

Figure 2. Binocular manipulation of White's effect. Binocular disparity is used so that the patches coaxial with the black stripes are at the same depth as are the black stripes and the patches coaxial with the white stripes are at the same depth as are the white stripes (both in front). Crossed fusers should fuse the two leftmost displays, and uncrossed fusers the two rightmost displays. White's effect should still hold.

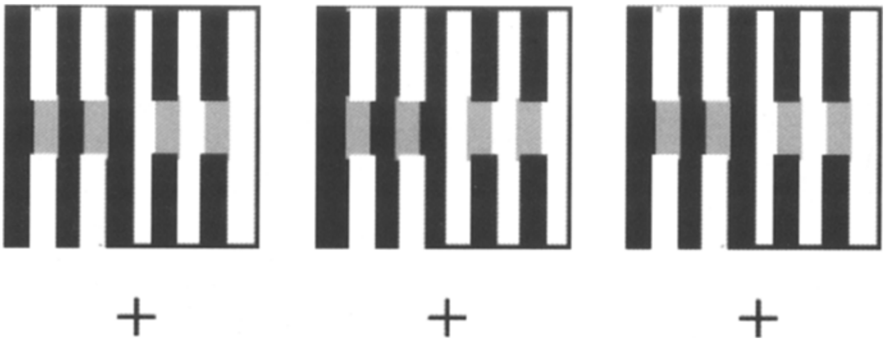

Figure 3. Binocular manipulation of White's effect. Binocular disparity is used so that the patches coaxial with the black stripes are at the same depth as are the white stripes and the patches coaxial with the white stripes are at the same depth as are the black stripes (both in back). Crossed fusers should fuse the two leftmost displays, and uncrossed fusers the two rightmost displays. White's effect should be decreased or even reversed.

In this display, the gray patches on the black stripes are perceived at the same depth as the black stripes (both together in front), and the patches on the white stripes are perceived at the same depth as the white stripes (again, both in front). This scenic grouping induced by depth is equivalent to the segmentation proposed for the flat display. We thus predict that perceived lightness for this depthful stimulus and for the standard White effect should be similar. This can be confirmed informally by fusing the stereo pair in Figure 2. A similar stimulus and result were recently reported by Spehar, Gilchrist, and Arend (1995).

Our hypothesis makes the prediction that modifying the depth relationships in White's effect so that inconsistent scenic segmentations are generated should alter perceived lightness. We were able to test this by again manipulating depth in such a way that the gray patches appear at the same depth as the noncoaxial stripes. Given this inconsistent segmentation, we would predict that White's effect will be decreased or even reversed. The stereo images shown in Figure 3 produce these depth groupings and show that the effect indeed diminishes (for some display configurations, it even reverses; see also Spehar et al., 1995). These informal results are currently being assessed psychophysically and have been confirmed by pilot studies.

It may at first seem counterintuitive to expect context segmentation effects in simple two-dimensional (2-D) scenes, such as the flat version of White's display; however, the contextual groupings and separations that we have invoked for it are consistent with simple T-junction cues to edge occlusion. Figure 4A shows the T-junctions and the occluding boundaries that they imply in White's stimulus. In recent work, Todorović (1997) has proposed that junction information (such as T-junctions and X-junctions) plays a critical role in the determination of lightness (see below).

SELECTIVE INTEGRATION MODEL

The light intensity at a point in the image is the product of surface reflectance and the illumination arriving at that point. Edge integration models propose that perceived lightness correlates with surface reflectance and employ three main computational stages in order to unconfound or disentangle the contributions of reflectance and illumination to image luminances: differentiation, thresholding, and integration stages. Differentiation plus thresholding is used as a means of *discounting the illuminant*. A properly chosen threshold is capable of eliminating all shallow intensity gradients within the image. For simple scenes, the shallow illumination component will thus be eliminated. At the same time, the sharp reflectance transitions will produce strong enough signals and be registered. Finally, differences are converted into lightnesses by integration operations in the final stage of processing.

Past edge integration schemes can be summarized by the following stages: $D \rightarrow T \rightarrow I$, where D, T, and I stand for the differentiation, thresholding, and integration stages, respectively (Horn, 1974). SIM builds on this scheme but suggests that integration must be *selective* in order to take into account scenic contexts and groupings. Therefore, within SIM, edge integration only occurs fully within scenic groupings and is attenuated between them. In this

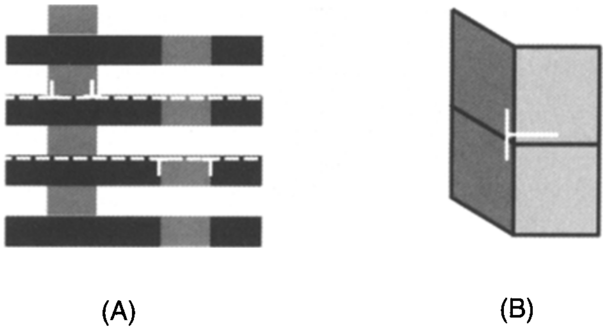

Figure 4. T-junctions. (A) T-junction cues to occluding boundaries separate White's stimulus into horizontal strips. (B) T-junction cues in more complex three-dimensional scenes, such as this folded card, separate surfaces.

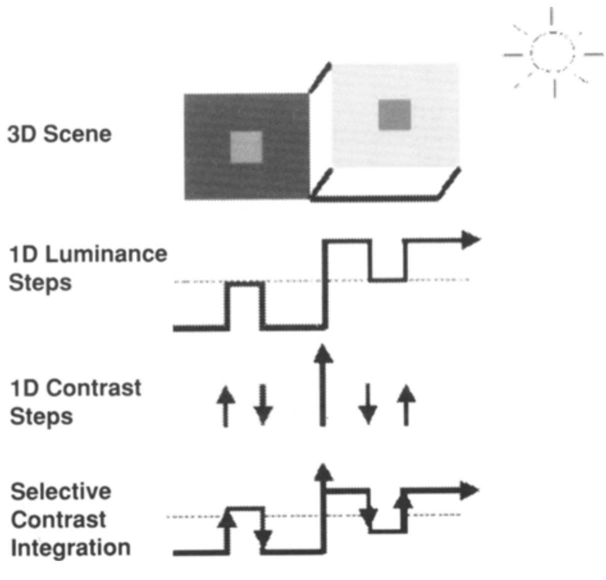

Figure 5. A one-dimensional cut through the scene on top gives rise to the luminance distribution shown (note that the space between the left block and the shadow to its right is only included for figure clarity; the two blocks can be considered to be abutting each other, or the left block touching the shadow). Selective integration correctly keeps the two blocks (lighting contexts) partially independent of each other. The gray patch on the right is determined to be darker than the patch on the right. Standard integration has no provision for separating lighting contexts.

way, edge information is not mixed indiscriminately across the scene. This allows the model to utilize initial contrast measures in a context-sensitive manner—as indicated by, for example, Gilchrist's coplanar ratio hypothesis.

The integration scheme in SIM can be stated as $C \rightarrow S \rightarrow I^*$, where C, S, and I^* stand for contrast measurement, contrast selection, and contrast integration stages, respectively. The contrast measurement stage substitutes for the differentiation stage and generates a measure of luminance ratios as the initial representation. Contrast ratios, as opposed to differences, discount not only illumination gradients, but also overall levels of illumination, allowing the model to achieve lightness constancy over a wide range of overall lighting conditions, in agreement with both the neurophysiological and the psychophysical data. Contrast selection occurs prior to contrast integration and achieves a selective reduction of those retinal contrasts between scenic context groupings, relative to those within context groupings. Context groupings in monocular 2-D images are determined by T-junction detection and processing.

To appreciate the difference between classical Horn integration and selective integration consider Figure 5. The two blocks define two separately lighted contexts that should remain partially independent during lightness determination. Therefore, selective integration emphasizes the small contrast steps within the blocks and partially suppresses the lighting contrast between the blocks. In this case, the gray patch on the right block is correctly determined to be darker than the patch on the left block (although it is more luminous).

Standard integration would, instead, fully integrate the large contrast step in the middle of the scene, which is associated with the depth and, thus, the lighting jump from one block to the other. By this method, the lightness of the whole right block would be "lifted up" (or equivalently, the lightnesses of the left block would be "pushed down"), causing the right central patch to be incorrectly predicted to be lighter than the left central patch. The two contexts should remain largely independent, as in selective integration, to produce the correct appearance.

In our model, selective edge integration I^* is instantiated by *lateral integration* mechanisms in which neighboring sites (corresponding to adjacent image points) communicate with each other. As will be formalized below, signals within the lateral integration stage communicate from neighbor to neighbor, so that the lightness ratio between any two regions, even regions that are far apart in image space, tends to be consistent with an integration of *all* the edge contrasts between those regions. Selected contrast measures provide constraints on local lightness relationships, and the determination of lightness can be seen as a constraint satisfaction process.

Figure 6 presents a schematic representation of SIM. The input luminance is initially coded by contrast measures. These are then used to generate context boundaries

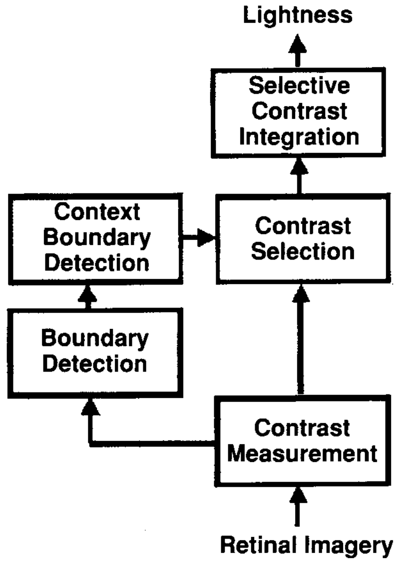

Figure 6. Macroscopic diagram of the selective integration model. Contrast signals are input to boundary detection and integration stages. Edge integration is regulated by context boundaries, where integration is reduced. The activity within the integration stage is the model's correlate of perceived lightness.

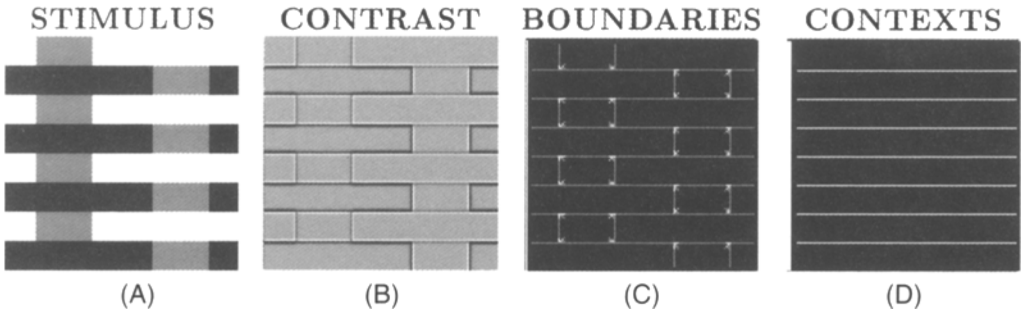

Figure 7. Computational stages of the selective integration model. (A) White's stimulus. (B) Stage 1 *on/off* contrast measures. (C) Stage 2 boundary signals. (D) Stage 3 context boundaries.

that divide up the image into context groupings; contrast measures are also used as input to the selective integration stage. Selective integration is modulated at context boundaries through the reduction of contrast measures at these boundaries.

Model Description

SIM consists of three processing stages, each of which consists of 2-D fields (or grids) of processing units, or *cells*. The input is also encoded in a 2-D field of activity. All connections between model stages are topographically organized in such a way that a spatial location (i,j) at a given stage connects to locations at or nearby (i,j) in its target fields. The activation level at individual model stages represents the output values of the respective stages. We describe the model with reference to simulations of White's effect and the Benary cross.

Input

The input images were prepared to match the relative luminance values recorded for real stimuli. Values used in simulations were based on photometer readings taken for our stimuli displayed on a CRT. Thus, typical white values where 30 times as large as typical black values, with middle gray typically around a third as large as the scenic white. Since the initial stage of our system codes local contrast ratios, the system is largely insensitive to the absolute magnitudes of the luminances resulting from white, gray, and black surfaces. For White's effect, we used the somewhat arbitrary input values of white = 300, black = 10, and middle gray = 100.

Stage 1: Contrast Measurement

At this stage, an *on* field with center–surround antagonistic cells measures the degree of local luminance contrast in input images coding stimuli luminances. The *on* field implements lateral inhibitory interactions that process the input luminance distribution. In uniform regions, a contrast measurement of zero results. Figure 7B shows the output of the *on* field for White's effect. In this figure, middle gray codes zero activity, and brighter and darker shades code activities greater and smaller than zero, respectively. The latter make up the signal of a corresponding *off* field.

Contrast responses are given by

$$X_{ij}^{(s)} = \frac{I_{ij} - S_{ij}}{K + S_{ij}}, \tag{1}$$

where I_{ij} is the input luminance (i and j denote spatial indices denoting the cell location on a two-dimensional lattice), S_{ij} consists of *surround* signals, and (s) denotes the spatial scale of the contrast response. More specifically, S_{ij} amounts to a measure of the average luminance in the vicinity of position (i,j)—given, for instance, by having a Gaussian weighting function modulate the I_{ij} contributions as a function of distance. This contrast equation includes a difference in the numerator but also a denominator including S_{ij}, which gives the system sensitivity over a wider dynamic range of overall lighting intensities. For example, the outputs of this equation are quantized to two orders of magnitude without loss of sensitivity to a similar range of luminance ratios, despite variation over eight orders of magnitude of overall input intensities. The small constant K in the denominator acts to suppress response to center–surround differences that are small, thereby making the system more resistant to noise.

Our scheme for contrast measurement is similar to the proposals to use difference-of-Gaussian operators to model the receptive field structure of retinal ganglion cells (Enroth-Cugell & Robson, 1966; Rodieck, 1965). In the classical proposal, the center and surround contributions combine linearly to determine the cell's response. The present approach adopts a multiplicative, or shunting, formalism (Furman, 1965; Grossberg, 1970; Hodgkin, 1964; Sperling, 1970; Sperling & Sondhi, 1968), in which center and surround contributions generate a nonlinear, normalized-difference measure proportional to the luminance ratios at borders.

In the present implementation, two spatial scales, $s = [1,2]$, of Stage 1 cells were employed. Contrast responses were computed for both a "small" and a "large" scale of

Table 1
Selective Integration Model Parameters for All Simulations

Name	Description	Value	Equation(s)
	Stage 1: *On* and *Off* Center–Surround Processing		
σ_{c0}	center blurring constant	0.3	??
σ_{s0}	surround blurring constant	0.82	??
σ_{c1}	center blurring constant	0.82	??
σ_{s1}	surround blurring constant	3.82	??
	Stage 2: Contrast Selection		
σ_w	width blurring constant	0.5	4
σ_l	length blurring constant	1.5	4
σ_{cs}	contrast selection blurring	1.0	??
σ_l^{cb}	context boundary contrast blurring	50.0	??
σ_w^{cb}	context boundary contrast blurring	1.0	??
P^2	contrast suppression larger-scale	0.75	??
P^1	contrast suppression smaller-scale	0	??

surround inhibition (given by Gaussian standard deviations σ_{s0} and σ_{s1}); the center contribution for the smaller spatial scale was the size of a single pixel, and the center contribution for the larger scale was determined by a Gaussian blurring the size of the surround for the smaller scale. Table 1 specifies the major model parameters.

For each scale s, both the X and $-X$ are rectified, breaking the information into two nonnegative channels called the *on* and *off* channels, respectively, which drive boundary detection (Stage 2a).

Stage 2: Contrast Selection

The contrast selection stage consists of three substages. The first of these, Stage 2a, detects oriented boundaries based on Stage 1 contrast measures. The second, Stage 2b, operates on this oriented boundary image map to detect context boundaries that segment the scene into context groupings. The third, Stage 2c, partially suppresses contrasts near context boundaries, fully selecting only those contrasts away from context boundaries.

Stage 2a: Boundary detection. This stage performs oriented boundary detection by employing elongated receptive fields such as those found in simple and complex cortical cells. The small scale *on* and *off* Stage 1 responses are used to guide the boundary detection stage. Our strategy follows that initially suggested by Marr and Hildreth (1980), who noted that if an *on* response occurs at position P and an *off* response occurs at a nearby position Q, a contour (or a zero-crossing in Marr & Hildreth's terminology) must lie somewhere in between P and Q. Stage 2 cells exhibit oriented odd–symmetric (bilobed) receptive fields that are excited by the oriented sampling of *on* and *off* Stage 1 cells. Responses are maximal when *on* activation is strong on one side of a cell's receptive field and *off* activation is strong on the opposite side. In other words, the cells are tuned to *on*/*off* contrast cooccurrence, or juxtaposition, as in previous models (Grossberg, Mingolla, & Williamson, 1995; Marr, 1982; Pessoa, Mingolla, & Arend, 1996). Both light–dark and dark–light oriented cells are employed.

The above computations are formalized as follows. First, *on* and *off* Stage 1 responses are convolved with oriented, offset Gaussians, producing boundary activations for the *on* (b^+) and *off* (b^-) channels:

$$b^{+k} = X^+ * G^k \tag{2}$$

and

$$b^{-k} = X^- * G^k, \tag{3}$$

where $*$ indicates a discrete spatial convolution, or filtering, operator, and k indicates the orientation of G^k, an oriented contrast-sensitive filter based on an oriented Gaussian kernel defined as follows:

$$G_{pq}^{lw} = \frac{1}{\sigma_l \sigma_w \sqrt{2\pi}} \exp\left[-.5\left(\frac{p^2}{\sigma_l^2} + \frac{q^2}{\sigma_w^2}\right)\right], \tag{4}$$

where (p,q) are spatial location indices and l and w define the spatial extent (length and width) of the operator (through σ_l and σ_w, respectively). For simplicity, in these simulations, we used only vertical and horizontal orientations, detecting boundaries of both directions of contrast or polarity. This yields a total of four boundary cell types b^k. For $k = 0,2$, orientation is vertical; for $k = 1,3$, the orientation is horizontal. G_{pq}^{lw} above is rotated to compose simple cells of different orientations and offset to sample *on* or *off* Stage 1 activity adjacent to the center of the receptive field.

In order to detect contrast changes, cells are tuned to *on*/*off* response juxtaposition, or adjacency. Thus, for a boundary detector of orientation $k = 0$, which is vertical with left-to-right polarity on-to-off, the activation at location (i,j), is

$$B_{ij}^k = f(b_{i-1,j}^{+k} \times b_{i+1,j}^{-k}), \tag{5}$$

with the nonlinear function f specified by $f(w) = 1$ for $w > 0$ and $f(w) = 0$ otherwise.

The activities of cells tuned to the same boundary orientation but opposite contrast polarities are combined, reducing the total number of boundary cell types from four to two (horizontal and vertical). In real scenes, additional orientations might be necessary for detecting diagonal boundaries. In the simple scenes we investigate, diagonal context boundaries are not present. Gaps in boundaries were completed using simplified rule-based BCS grouping processes (Grossberg & Mingolla, 1985a, 1985b; Grossberg, Mingolla, & Ross, 1997). Figure 7C shows the final Stage 2 boundary activations summed across the two orientations employed.

Stage 2b: Context boundary detection. The central proposal of this paper is that context information modulates the process of contrast integration for the determination of lightness. The problem of scene segmentation based on depth, motion, configuration, and other cues is perhaps the most outstanding problem in vision research. Our aim here is not to attempt to solve this problem but to focus on how segmentation *interacts* with lightness perception. Accordingly, we adopt a simplified scheme in which T-junction information is employed for the determination of context boundaries. The T-junction scheme

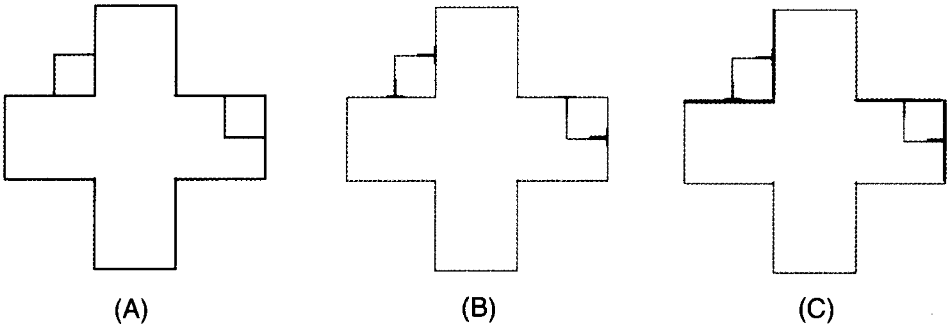

Figure 8. Determination of context boundaries. (A) Outline of a Benary cross. (B) T-junctions detected. (C) T-junction activation is used to define occlusion or context boundaries. T-junction activation spreads along the boundary responses of Stage 2a, starting from the occluding crossbars of the Ts.

is sufficient for recovering context boundaries for several stimuli simulated in this paper, but it is not intended as a substitute for a full model of segmentation, which should include more sophisticated binocular and monocular mechanisms. Moreover, what we call *T-junction detectors* in our simulations are not specified to exclude responses to more complex junctions that include T boundary formations. In other words, our T-junction detectors sometimes also respond strongly to the presence of X-junctions and Psi-junctions. We found these simple detectors to be sufficient for simulating the data that we address in this paper. However, we expect that the visual system is more discriminating than our model and does distinguish among junction types in detecting scenic structure.

The detection of context boundaries by T-junction detection involves two processes. First, T-junctions of various orientations are detected across the image. Next, these isolated junctions, in conjunction with each other and with boundary information, are employed to delineate entire context boundaries. These two processes are best illustrated through an example. Consider the Benary cross outline stimulus (Figure 8A). First, T-junctions of various orientations are located within the image.

Initially, dedicated cells detect T-junction configurations in the image. For example, one type detects the T-junction where left, right, and top positions are active at Stage 2 (leftmost T-junction in panel B). Corresponding cells detect T-junctions of different orientations. The "horizontal" context boundaries signaled by upright or upside down T-junctions can be detected by cells tuned to the following summation of T-like spatial boundary arrangements:

$$T^{\mathrm{h}}_{ij} = f(B^{\mathrm{h}}_{i-d,j} \times B^{\mathrm{h}}_{i+d,j} \times B^{\mathrm{v}}_{i,j+d}) + f(B^{\mathrm{h}}_{i-d,j} \times B^{\mathrm{h}}_{i+d,j} \times B^{\mathrm{v}}_{i,j-d}), \tag{6}$$

where d is a constant spatial displacement, the index h indicates horizontal orientation, the index v indicates vertical orientation, and f is defined as above. The "vertical" context boundaries signaled by lying down T-junctions are detected by cells tuned to the following summation of T-like spatial boundary arrangements:

$$T^{\mathrm{v}}_{ij} = f(B^{\mathrm{v}}_{i,j-d} \times B^{\mathrm{v}}_{i,j+d} \times B^{\mathrm{h}}_{i+d,j}) + f(B^{\mathrm{v}}_{i,j-d} \times B^{\mathrm{v}}_{i,j+d} \times B^{\mathrm{h}}_{i-d,j}). \tag{7}$$

Then, to recover completed context boundaries given the initial T-junction responses, the T-junction activations T^{h} and T^{v} were allowed to spatially spread along contiguous detected boundaries B^{h} and B^{v}, respectively. This spreading allowed the delineation of entire context boundaries y^{h} and y^{v}. For the Benary cross, a pictorial illustration of the boundaries marked by this T-junction-based context boundary detection method is shown in Figure 8C.

In some images, contradictory context segmentations are supported by different T-junction activations. To resolve such conflicts, intersecting context boundary activations are allowed to compete at the sites of intersection. Simple rules determine which context boundary y survives such competitions to determine the final deconflicted context boundary maps Y^k; The context boundary supported by the most T-junctions wins such competitions, with ties being broken according to the strength of the underlying boundary activations B, themselves proportional to the strength of the contrast across the boundary. Once again, it is not our intent to claim much for our T-junction scheme as a means of context segmentation.

One implication of our model, as well as others like it, is that lightness effects themselves may prove useful in probing how the visual system combines cues to accomplish segmentation. Figure 7 shows the final Stage 3 context boundaries, Y^k, generated by the model for White's effect.

Stage 2c: Selective contrast suppression. In Stage 2c, the context boundaries Y^k found in Stage 2b are used to selectively suppress those Stage 1 contrast measures associated with them. Contrast suppression is achieved through a three-step process: (1) selection of those contrasts that should be partially suppressed, (2) determination of the average contrast across selected context boundaries, and (3) partial suppression of the average contrast across the selected context boundaries.

1. Initially, contrast measures of Stage 1 on either side of a context boundary (Stage 2b) are selected. Mechanistically, this selection was achieved through the multiplicative gating of contrast measures X with a Gaussian blurred context boundary map Y^k to produce W^k,

$$W^k = X \times Y^k * G_{cs}, \quad (8)$$

where $X = X^+ - X^-$ are the combined Stage 1 contrast responses, Y^k are context boundaries for orientation k, and G_{cs} is an isotropic spatial-blurring Gaussian.

2. Next, the selected contrast measures W^k are blurred along the context boundary to obtain a measure of the average contrast across the context boundary. Oriented Gaussian blurring in parallel to the context boundary measures the average contrast across the context boundary and propagates this average measure all along the length of the context boundary:

$$V = W^h * G_{cb}^h + W^v * G_{cb}^v, \quad (9)$$

where G_{cb}^k is a Gaussian elongated in direction k.

3. Finally, a proportion P of the selected average contrast measures V were subtracted from the large-scale Stage 1 contrast measures. This subtraction, or partial suppression, of the Stage 1 contrast measures effectively reduces the constraints on contrast across the context boundary. In other words, the suppressed contrast measures are taken into account but now are associated with smaller lightness differences. Formally,

$$Z^{(2)} = X^{(2)} - P^{(2)} \times V \quad (10)$$

and

$$Z^{(1)} = X^{(1)}, \quad (11)$$

where the superscripts (1) and (2) denote the respective spatial scales. The resulting contrast measures, Z, are then integrated to model lightness measures (Equation 12). In our simulations, we used $P^{(2)} = .75$ and found that this suppression of the larger scale contrasts were sufficient to model the effects. We used $P^1 = 0$ for simplicity; larger P values give larger effects of context on contrast; smaller P values reduce the effects of context on contrast.

Stage 3: Selective Contrast Integration

Integration of the selected contrast measures Z was achieved through the iteration of a lateral, multiscale integration equation. Discussion of how iteration of such a contrast inversion can eventually yield stable lightness measures can be found in Horn (1974) for the special case of difference contrast measures. Our proposal works in a similar way, but for the iterative inversion of contrast *ratios*. Simply put, iteration of the equation for l below spreads the constraints X until the whole field of activations l tends to satisfy the full set of local contrast ratio measures given by X:

$$l = (1 + Z^{(1)}) \times (1 + Z^{(2)}) \times (K + Sl), \quad (12)$$

where Sl denotes a Gaussian blurred version of the time-varying lightness map l (in the previous iteration) and K is the same constant as that used for Stage 1. The main idea of Equation 12 is that it attempts to recover lightness values from the array of contrast signals X by approximately inverting the contrast measurement equation (see Equation 1). It should be noted that an exact inversion at a single scale would take the form $l = X \times (K + S) + S$. To simplify computations, our equation combines integrations across the two scales and also adds an additional small amount K to the large-scale integration on each iteration.

A similar multiscale integration technique is used in the reconstruction of images from a Laplacian pyramid (Burt & Adelson, 1983). The interscale redundancy exploited both in Laplacian reconstruction and in our own model is that integrated large-scale intensities can serve as reconstructed surrounds for the smaller scale.[2]

The added integration constant K has the effect that all lightness values tend more toward the maximum value. In the limit, it means that, at equilibrium, l tends toward white and that the maximum lightness value in any given scene also tends toward white. This tendency is in agreement with *highest luminance anchoring rule* that Gilchrist and colleagues have recently investigated.

In simulation, initial conditions were set such that all values of l were equal to the maximum (white) $l = 1.0$. During integration, l values were bounded between 1.0 and .03. This ensured an upper limit of white and a lower limit of black. Further work would be necessary to model the perception of *self-luminous* or *black hole* regions that would fall outside of this range.

Lightness

The final lightness values L associated with the lightness percept were given by iterating Equation 12 a fixed number of times and then taking the square root. The square root transformation rendered our final values on a scale similar to the Munsell Value scale, thus permitting a more direct comparison with published data. The final model output was given by

$$L = 10 \times \sqrt{l}. \quad (13)$$

MODEL SIMULATIONS

White's Effect and the Benary Cross

From the standpoint of classical lateral inhibition proposals, White's effect presents a striking violation of simultaneous contrast since the contour length of the gray patches is larger for the stripes they are not coaxial with. Therefore, simultaneous contrast predicts the opposite percept of what occurs—namely, that for the display of Figure 1, the gray patches on the left should appear lighter than the gray patches on the right. As is clear from the display, just the opposite relationship is seen. In an earlier section, we advanced the notion that White's effect can be explained by context-selective integration mechanisms. In particular, a segmentation of the stimulus into horizontal stripes was assumed to provide the proper context for lightness determination.

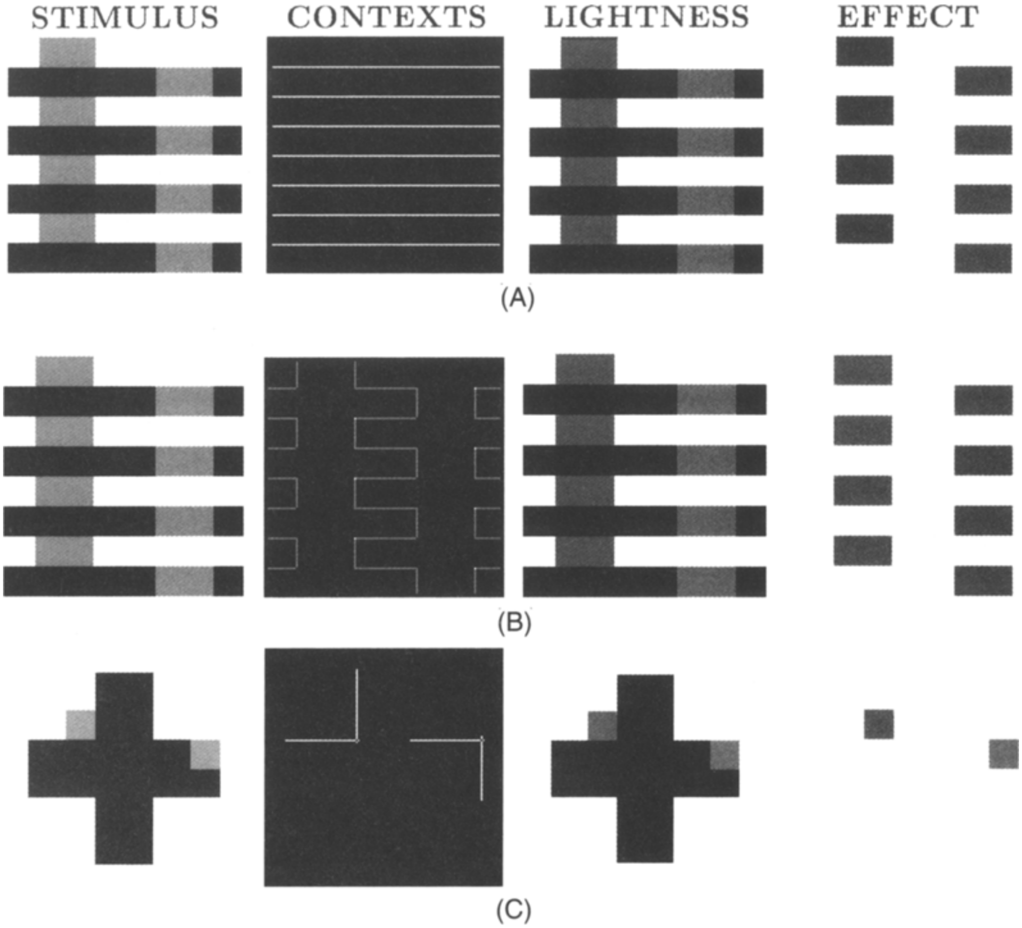

Figure 9. Simulation of White's effect and the Benary cross by the selective integration model. Detection of stimulus (first column) T-junctions allows the identification of context boundaries (second column). Context boundaries modulate a selective integration process that determines scenic lightnesses (third column). Isolating the test patches (fourth column) shows the effect. (A) For White's effect, simulation resulted in lightnesses of 4.8 for the leftmost patches and 5.7 for the rightmost patches (white = 9.8 and black = 3.0). (B) For White's effect with the depth organization conflicting the segmentation given by T-junctions, the opposite relationship is found, with the leftmost patches at 5.4 and the rightmost patches at 4.8. Note that externally provided context signals indicate that the left patches are grouped with the black bars (including the rightmost endings) and the right patches are grouped with the white bars (including the leftmost endings). (C) For the Benary cross, simulation resulted in lightnesses of 4.4 for the leftmost patch and 5.6 for the rightmost patch.

This explanation of White's effect has been supported through computer simulations (Figure 9). Like lateral inhibition proposals, SIM initially computes contrast ratios. Lightness is not a direct correlate of contrast measures (as in previous filling-in schemes; e.g., Grossberg & Todorović, 1988) but, instead, is determined by a selective integration process that strongly takes into account scenic groupings or contexts. For White's display, contexts consist of horizontal regions including black or white bars, together with the associated gray patches. These are produced by SIM's T-junction-guided process. For White's display, although two types of T-junction occur—one indicating that the black stripes occlude the white stripes, the other indicating the converse arrangement (see Figure 4)—both are compatible with the parsing of the scene into horizontal contexts. Hence, the gray patches coaxial with the black bars have their lightness determined in relation to the black bars, and the gray patches coaxial with the white bars have their lightness determined in relation to the white bars. The noncoaxial regions still exert an effect on the adjacent region, but the effect is mitigated owing to the reduction of integration—thus, for example, the white regions's effect on a gray patch coaxial with the black stripes is weak. The result is that patches parallel to the black bars are lighter than the patches parallel to the white bars. Essential to this result is, of course, the proper determination of scenic segmentations.

Our context-based account of White's effect motivated us to attempt to modify the assumed scenic grouping through the use of disparity cues. We thus investigated the two binocular manipulations of White's effect (Figures 2 and 3). To test the model, we presented it with what we previously called the inconsistent version of White's effect (Figure 9B). Context boundaries are used by SIM to guide the selective integration of lightness signals. In general, context boundaries are indicated by any of a large number of scenic cues, such as motion and depth, as well as by configurational cues, such as T-junctions. The present simulation assumes that externally provided context signals, corresponding to the depth organization of the scene, provide the appropriate scenic segmentation (see the Discussion section). The depth cues of the inconsistent stimulus are such that the gray patches on the left are grouped with the black bars and the gray patches on the right are grouped with the white bars—the opposite of the grouping generated by the T-junction segmentation process for the standard version of White's effect. This segmentation implies that contrast information associated with the white bars will strongly affect the appearance of the left patches and contrast information related to the white regions will strongly affect the right patches. SIM therefore produces a reversal of White's effect, with the gray patches on the left having a higher lightness than the ones on the right.

The Benary cross constitutes another example of a challenging stimulus from the standpoint of simultaneous contrast accounts (see Figure 9C). Although the two patches exhibit equivalent contours, the gray patch on the left appears darker than the one on the right. Lateral inhibition (or contrast measures) cannot explain this effect. SIM's account of the Benary cross is similar to its explanation of White's effect. T-junction information is used to create two distinct contexts for the right and the left (same-luminance) gray patches. One context comprises the cross, including the right gray patch; another comprises the remaining scenic regions, including the left patch (see Figure 8). Therefore, the lightness of the left patch is determined mainly in relation to the high-luminance background, and the lightness of the right patch is determined mainly in relation to the low-luminance cross. A darker patch on the left, when compared with the one on the right, results.

Shading and Transparency Effects

T-junction information proved quite effective in determining scenic contexts for the simulations of the standard version of White's effect and the Benary cross. Figure 4 shows T-junctions associated with the stimulus for White's effect but also includes T-junction cues in more complex 3-D scenes, such as those for the folded-card or corner stimulus. Although the junctions in this stimulus are actually X-junctions, such configurations are also detected by our T-junction detectors since they include T-shaped boundary arrangements. As has been discussed above, our T-junction detectors also respond strongly to T-shaped boundary arrangements embedded in psi-junctions.

In the stimulus shown in Figure 10A, the patches in the third column in the second and fourth rows have the same luminance but appear quite distinct. Adelson's

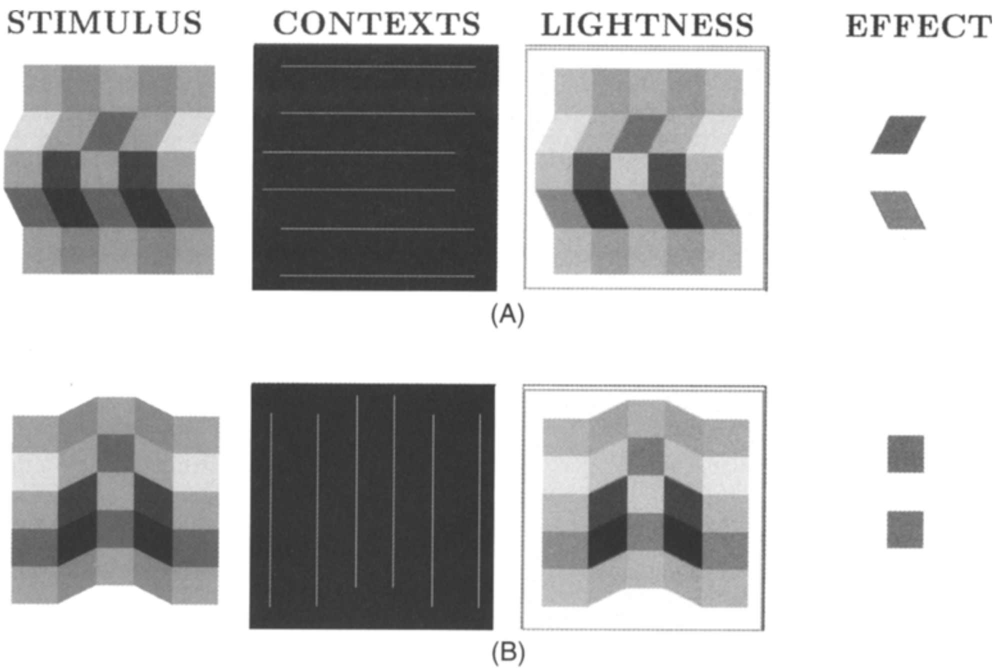

Figure 10. Simulation of Adelson's folded-card stimulus by the selective integration model. The folded card in panel A yields horizontal T-junction boundaries that separate the patches into different segmentation contexts. In panel B, T-junction boundaries group the patches together in the same vertical context. The magnitude of the lightness difference between the patches in panel A is greater (top patch = 4.1 and bottom patch = 4.5) than that in panel B (top patch = 4.19 and bottom patch = 4.15).

(1993) original explanation for this difference is related to the perception of different levels of illumination falling on the different planes. Thus, the top patch is seen as a dark gray patch that is brightly lit, whereas the bottom patch is seen as a light gray patch that is dimly lit. This type of explanation seemed reasonable, given a much smaller lightness difference in a rotated version of the stimulus (Figure 10B), in which the two patches are perceived as lying in the same plane with the same illumination.

T-junctions can be used to provide segmentation contexts in the Adelson stimuli. Moreover, these are consistent with the shading contexts assumed by Adelson. In Figure 10A, T-junctions occur in such a way that the crossbar of the Ts are horizontal. If these T-junctions are used to mark the contours that are collinear with these crossbars, the card is segmented into horizontal strips (see the section on Stage 2b: Context Boundary Detection), each providing a separate *shading* context. Within each segmentation context, our model integrates the effects of local contrasts to yield lightness measures. However, across context boundaries, the influence of contrast measures is reduced. Thus, in Figure 10A, the lightnesses of the top and bottom patches of interest are determined primarily on the basis of their contrast with other patches within the same row. As a result, the upper test patch, which has the lowest luminance in its row, appears significantly darker than the lower test patch, which has a high luminance in its row. Accordingly, a similar analysis can be performed for the stimulus in Figure 10B. In this case, T-junctions indicate columnar contexts that place both test patches within the same segmentation context, which is consistent with a single shading context. Therefore, their lightness is primarily determined by relations to the same set of patches. However, their horizontal neighbors also have an effect on their appearance—albeit, a reduced one, given the existence of interposing context boundaries. In all, a smaller lightness difference is produced for the top and bottom patches in this arrangement, when compared with the configuration of Figure 10.

The stimuli that we have considered up to now have been explained as the result of the modulation (reduction) of the influence of *local* contrast measures at context boundaries, as the diminished effect of the white background on the adjacent right patch in the Benary cross. Although we have so far stressed these local interactions, the same processes lead sometimes to nonlocal effects as they reorganize the lightness assignments within a scene. This is illustrated by considering another stimulus proposed by Adelson (1993) as an example of the effect of transparency on perceived lightness. In Figure 11A, the top and bottom rows of diamonds have the same physical gray level, although they appear to have different lightnesses. This difference was originally considered by Adelson to be the result of the perception of transparency; the bottom diamonds are thus seen as made of a light gray material that is covered by a transparent film. As evidence for his transparency explanation, Adelson presented the stimulus shown in Figure 11B, in which the top and bottom rows of diamonds are the same as those in Figure 11A. Note that the difference in lightness between the dia-

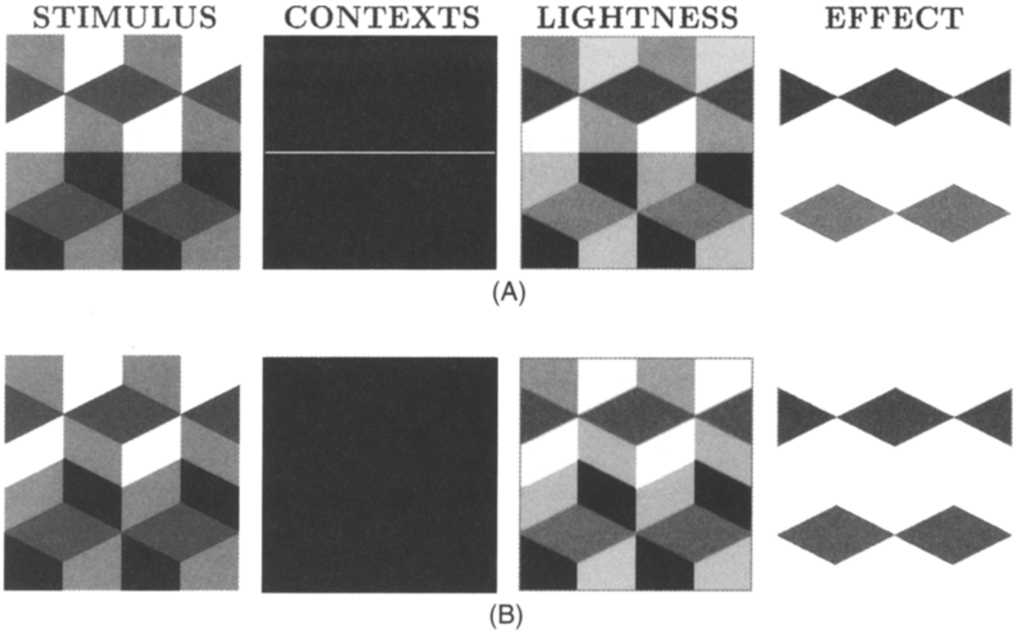

Figure 11. Simulation of Adelson's transparency effects by the selection integration model. T-junctions divide the scene into top and bottom transparency contexts in panel A but not in panel B. The top diamonds in panel A resulted in a lightness of 4.0; the bottom diamonds gave a lightness of 4.5. In panel B, the top diamonds had a lightness value of 4.1, and the bottom diamonds one of 4.15.

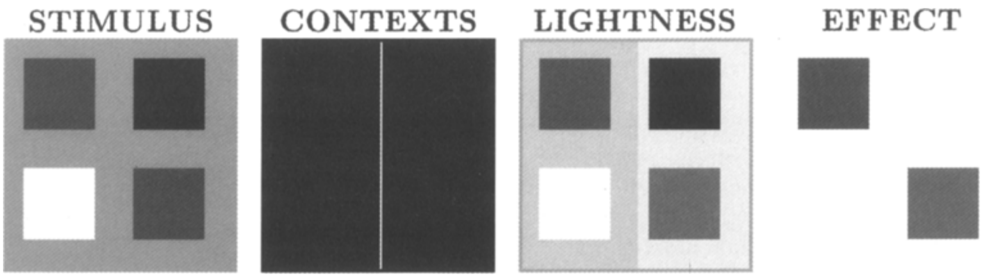

Figure 12. Simulation of the role of belongingness in lightness perception, as studied by Agostini and Proffitt (1993) in their (motion) common fate paradigm. If a motion context boundary is added to this "zoomed in" version of the stimulus, so that the region is divided in half vertically, the upper left and the lower right gray patches exhibit lightnesses of 4.0 and 4.5, respectively.

monds in the top and bottom rows for this stimulus is greatly reduced.

According to our model, a key difference between the above scenes is that in the first one (Figure 11A), T-junctions can be used to separate the top half of the scene from the bottom half. This segmentation effectively creates two domains of influence (top and bottom) within which contrast measures exert their maximal influence. Therefore, our model determines the lightness of the bottom row of diamonds, using the contrast measures of the relatively dark patches in the bottom half of this scene, resulting in a relative lightening of the bottom row of diamonds. At the same time, in the scene shown in Figure 11B, no such T-junction context boundary separates the top half from the bottom half of the scene. Consequently, the scene is not divided up into two domains of influence but, instead, is composed of a single one. As a result, all scenic contrast measures are taken more effectively into account in the determination of the appearance of the diamonds, which end up having similar lightnesses.

It is important to emphasize that the above explanation is nonlocal insofar as there are no context boundaries that directly abut the diamonds themselves. In the present case, it is the *integration process* that ends up being weakened across an occluding boundary (Figure 11A), resulting in nonlocal effects. It is worth noting that although we have shown how our model can account for a lightness percept in a stimulus consistent with the occurrence of transparency, we are not claiming that transparency is not perceived or that it has no effect on lightness perception. We only mean to show that even in a simplified model in which scenic lightnesses are represented in a single layer, selective integration can account for such effects.

Perceptual Belongingness and Gestalt Explanations

Our proposal of selective integration is not unlike Gestalt explanations of the Benary cross and White's effect, which have relied on grouping or *perceptual belongingness* to determine lightness (Agostini & Proffitt, 1993; Benary, 1924; Rock, 1983; Wallach, 1948; Wolff, 1933). The appeal of Gestalt explanations notwithstanding, their utility remains limited, given the lack of operational definitions for such critical concepts as *perceptual belongingness*. Our model overcomes part of this deficiency—by showing how a simple type of T-junction scheme can help define scenic contexts, or groupings—while retaining what we think are the key insights of the Gestalt explanations.

Recently, Agostini and Proffitt (1993) presented an investigation of *belongingness* and lightness perception. These authors were able to demonstrate nonlocal effects of motion segmentation on perceived lightness. Common fate was used to group two collections of dots, one composed of black elements, the other of white elements. Two identical gray dots were also part of the display, one sharing the motion of the collection of black dots, the other sharing the motion of the white dots. Subjects judged the gray dot that moved with the black dots to be lighter than the dot that moved with the white dots. This difference disappeared when a stationary version of the display was used (in which no common fate existed to group scenic elements).

Figure 12 shows a simpler version of the Agostini and Proffitt (1993) stimulus used to probe SIM's behavior. It was assumed that motion cues provide the appropriate segmentation of the scene into two separate contexts, one containing a black element and a gray patch, the other containing a white element and a gray patch. As in the case of our simulations of static displays, the context boundary effectively isolates the two contexts guiding the domain of influence of the initial contrast measures. As in the Agostini and Proffitt study, the gray dot grouped with the black dot is lighter than the one grouped with the white dot.

Coplanarity and Surface Grouping

In a series of experiments, Gilchrist (1977, 1980) showed that depth configuration and spatial layout help specify the perception of lightness. More specifically, he proposed that the ratios of coplanar surfaces, and not necessarily the retinally adjacent ratios, determines light-

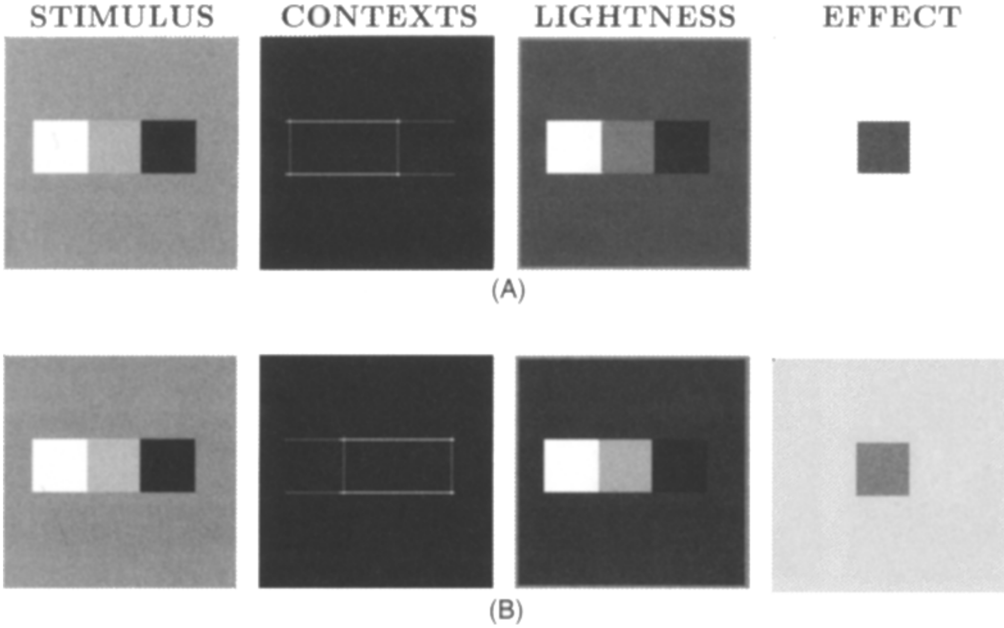

Figure 13. Simulation of Gilchrist's coplanarity experiment. When the test patch is grouped in the same context with the lighter inducing patch by coplanarity, as in panel A, the lightness is relatively dark at 5.0. However, when the patch is grouped with the darker inducing patch, as in panel B, its lightness is relatively light at 7.0.

ness—the so-called coplanar ratio hypothesis. Gilchrist was able to illustrate his hypothesis by comparing the lightness percepts resulting from two displays that were similar in terms of their 2-D configuration but gave rise to very different depth percepts. Other data have also been reported suggesting the weakening of contrast effects through the introduction of depth in disk–annulus stimuli (Gogel & Mershon, 1969), in the Koffka ring (Wist & Susen, 1973), and in Mach bands (Wist, 1974).

SIM is consistent with the above findings since our selective integration mechanisms will be affected by the depth manipulations in these displays in such a way that contrast measures associated with context (depth) boundaries will have a decreased effect on perceived lightness. Figure 13 shows how the model accounts for a schematic version of Gilchrist's coplanarity experiment. Scenic groupings directly defined by 3-D depth cues are provided externally and define display contexts. In one of the displays (Figure 13A), the gray and white squares belong to a front plane, with the black square being perceived in the back. In this case, the gray square will be darkened by simultaneous contrast. In the other display (Figure 13B), the gray and the black squares belong to the front plane, with the white square being perceived in the back. Selective contrast integration lightens the square. It should be noted that in these simulations, the actual (internal) array of luminance values produced a luminance range that exceeded the 30:1 ratio suggested as important for generating depth effects in such displays (Gilchrist, 1980).

Absolute Lightness and Anchoring

It is widely recognized that relative luminance ratios can only produce relative lightness values. However, as Arend (1973) states, "our perceptual scales are not interval but rather have an origin" (p. 341). This important issue goes back to Hering's (1874/1964) assignment of "middle gray" to the luminance that corresponds to the midpoint of the range of perceptual responses. Gilchrist and colleagues have recently investigated the rules for *anchoring*, or providing absolute lightness values. Their investigation of a *staircase* Gelb effect illustrates the key issue (Cataliotti & Gilchrist, 1995). In their study, the appearance of a given patch turns darker and darker as more and more higher reflectance patches are introduced into the scene. These drastic changes in appearance depend on physical changes occurring at distant regions in the scene—that is, they are nonlocal. From this and other studies, Gilchrist and colleagues have suggested that a process by which absolute lightness values are assigned is central to the determination of lightness.

Although we have not explicitly incorporated anchoring processes into the current version of SIM, our scheme handles the Cataliotti and Gilchrist (1995) finding appropriately. Figure 14A shows that when a single gray patch is seen on the display, it appears quite light. This occurs because it is the highest luminance in the scene and is not pushed down from its tendency to approach a white lightness value. As other higher luminance patches are introduced into the scene, the original patch appears

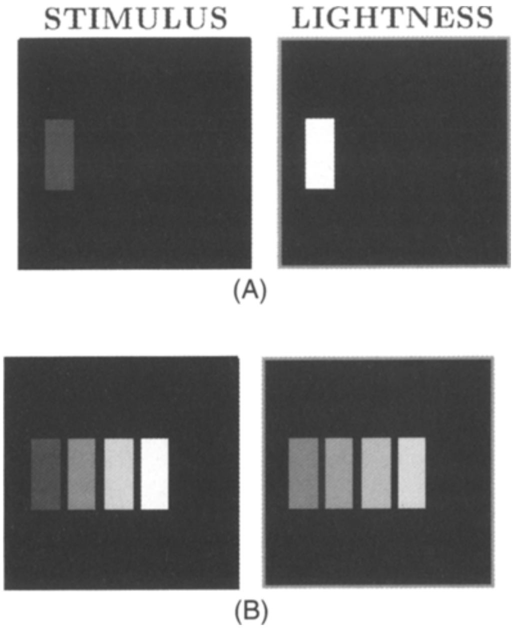

Figure 14. Simulation of the staircase Gelb effect. In the scene in panel A, the gray patch yielded a lightness of 10.0. In the bottom display (B), its lightness was reduced to 7.0.

much darker (Figure14B). This occurs as each new patch of a higher luminance will contribute to the darkening of the leftmost patch, through the integration of the contrast measures introduced. This is another example of SIM's ability to deal with nonlocal effects.

Contrast Effect and Haploscopic Display

Several researchers have proposed that when observers are asked to match lightness, two mechanisms oppose each other. Gilchrist (1988, 1994) has proposed that there is a compromise between ratio matching and luminance matching, as revealed through the analysis of the data of several investigations. In fact, Gilchrist (1988) proposes that the two mechanisms are in competition. Whittle and Challands (1969) and Whittle (1994) also proposed a similar scheme. A striking result in the Whittle and Challands study was that increments and decrements *never* appeared to be of the same lightness (i.e., could never be matched), regardless of magnitude. The haploscopic method employed by Whittle and Challands has been considered a paradigm that minimizes the competition between the two mechanisms (Gilchrist, 1994), thus providing an efficient paradigm for probing *ratio processing* in lightness perception. Gilchrist (1994) suggests that the way in which luminance matching interferes with ratio matching is through Land-type ratio *integration* (Land & McCann, 1971), where the ratios across several borders are integrated, as in the present model. According to this view, a more pure contrast effect occurs in the haploscopic displays simply because of the poverty of boundaries. Likewise, small deviations in lightness constancy in shadowed scenes where ratios are good lightness predictors can be viewed as the effect of incomplete exclusion of the illumination or shadow boundary from the edge integration process (Gilchrist, 1994).

Another way in which the haploscopic paradigm can be understood to isolate ratio processing is that it may provide a situation in which contrast integration is ineffective. The present proposal is that in the haploscopic display, binocular summation results in an equivalent background for both test and standard patches at the binocular level of processing, where integration would be assumed to occur selectively on the basis of depthful segmentations. Therefore, the baseline background level for both patches in the haploscopic display is identical, and the full contrast effect is preserved by contrast integration.

Figure 15 shows model simulations for binocular versions of both a standard simultaneous contrast display and a haploscopic stimulus. The contrast signals shown in B and F correspond to Stage 1 center–surround responses for the left- and right-eye images. Note that after binocular combination, the standard display includes a contour between the black and the white surrounds (C), whereas no analogous contour is present in the haploscopic version of the display (G). Accordingly, the difference between the lightnesses of the two gray patches is greater in the haploscopic display precisely because of the absence of a boundary between the two backgrounds—and the consequent lack of integration of contrast measures. In other words, the "pure" simultaneous contrast effect is revealed by the haploscopic configuration.

The above account of the haploscopic stimulus suggests the following prediction of the model. If the contour between the black and the white surrounds in a standard simultaneous contrast stimulus is stabilized, a stronger lightness difference should be revealed (when compared with a nonstabilized version). Qualitative evidence supporting this prediction of the model was obtained by Yarbus (1963), who revealed a much stronger effect in the stabilized condition. Evidence in this direction was also obtained by Piantanida and Gilchrist (see Gilchrist, 1994). Quantitative measurements of the difference between the stabilized and the nonstabilized versions would provide an indication of the extent to which integration processes "oppose" simultaneous lightness contrast.

DISCUSSION

T-Junctions and Grouping

Since we began the work presented here (Pessoa & Ross, 1995, 1996), several other interesting proposals for the importance of junction information in the determination of lightness have appeared. Todorović (1997) has proposed that junction information, such as that from T- and X-junctions, determines lightness in a way that may be described through *rules*. For instance, his T-junction rule

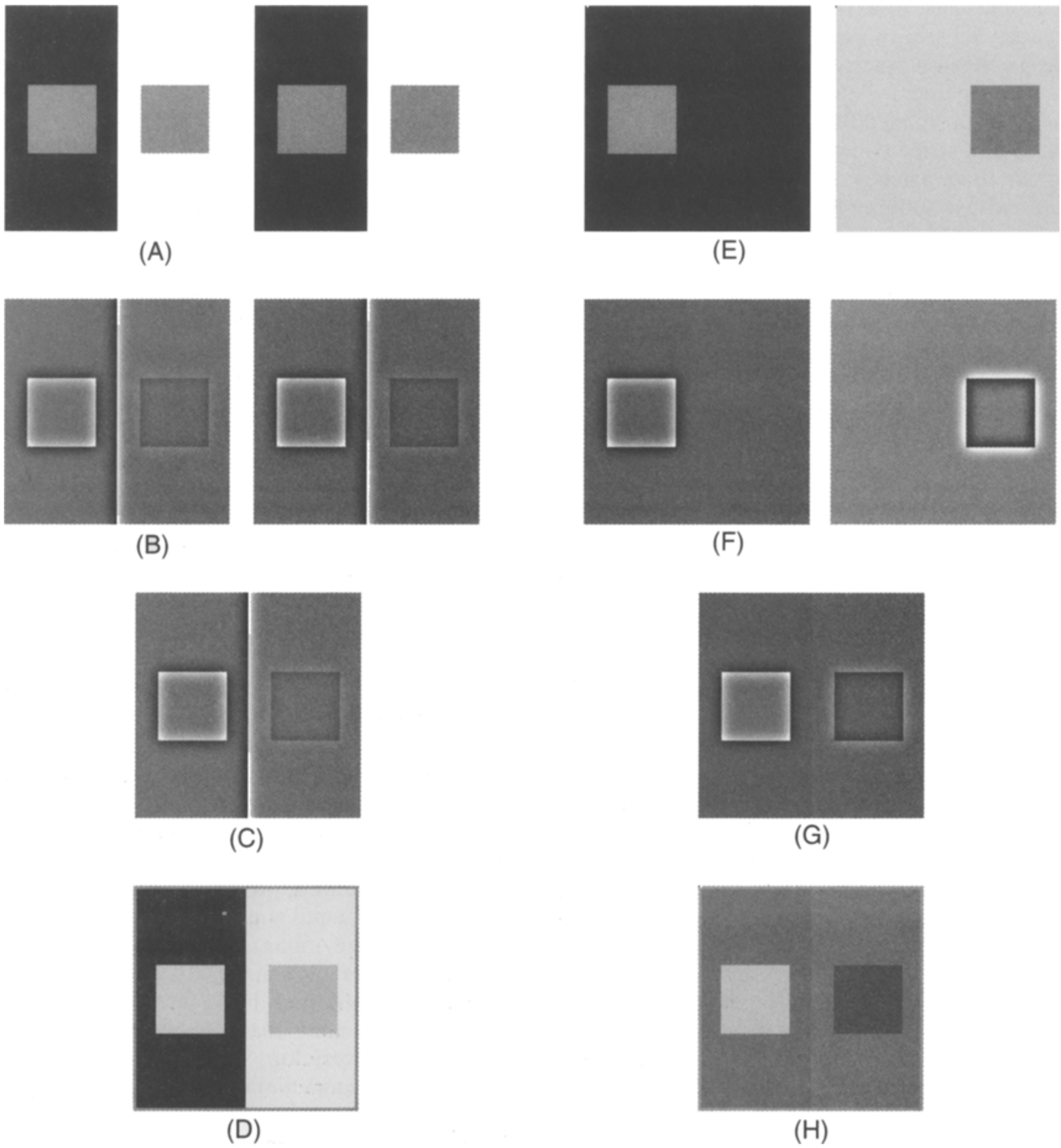

Figure 15. Simulation of binocular stimuli. (A) Left-eye and right-eye stimuli for the simultaneous contrast stimulus. (B) Retinal contrast for the simultaneous contrast stimulus. (C) Summed contrast measures. (D) Filled-in lightness for simultaneous contrast. The leftmost patch exhibited a lightness value of 5.3, and the rightmost patch a value of 5.0. (E) Left-eye and right-eye stimuli for the haploscopic display stimulus. (F) Retinal contrast for the haploscopic display stimulus. (G) Summed contrast measures. (H) Filled-in lightness for haploscopic display. The simultaneous contrast effect increased with the leftmost patch at 6.5 and with the rightmost patch at 4.7.

specifies that the lightness of patches that share edges with several other regions and whose corners involve T-junctions is predominantly dependent on the luminance of collinear regions—with the collinear regions consisting of the two regions that abut the stem of the T. In other words, the occluding region contributing to the T-junction does not significantly contribute to the process determining the lightness of the regions abutting the stem of the junction.

Anderson's (1997) proposal also relies heavily on T-junctions. His idea is that scission is a central component of the determination of surface appearance and that junction information is used to determine when scission occurs. T-junctions determine decomposition into layers so that the presence of a T-junction indicates the presence of overlying semitransparent layers. For White's effect, the patches on black are seen as lighter because some of their lightness is attributed to the black stripes, because of the postulated (unconscious) perception and subtraction of a stripe contribution to patch brightness. At the same time, patches on the white stripes are seen as darker because some of their brightness is attributed to the white stripes.

Our model shares with these two proposals the idea that junction information is important for lightness perception in some cases. Our emphasis, however, is on the role scenic segmentations have in generating selective contributions to lightness perception in general. White's effect is the result of T-junction information determining that the *horizontal contexts* should have a disproportionate effect on the lightness assignment of the gray patches, given potential occlusion relationships in the stimulus.

Our proposal suggests that other cues to the same scenic groupings, such as depth or motion, will generate the same lightness relationships. At the same time, cues that are able to induce new scenic groupings will alter the effect, by influencing which scenic contrasts are selectively integrated. Some specific predictions along these lines were informally tested in Figure 2 and are currently under further investigation.

Integration Mechanisms

As was initially discussed by Horn (1974), integration may be implemented in a number of ways. The question of how the visual system realizes such computations is still open. Although it is premature to discuss a clear mapping between cortical physiology and integration computations, we think that the experiments on the temporal dynamics of brightness by Paradiso and colleagues (e.g., Paradiso & Nakayama, 1991) provide a promising direction that could lead to bridging the gap between physiology and computational theory. In particular, we identify their temporal *filling-in* results with the lateral interactions achieved by contrast integration mechanisms (Stage 3 of the model).

Filling-in theories, as well as the present proposal, model lightness by explicit computations corresponding to the percept at each point in the stimulus array (see, e.g., Grossberg & Todorović, 1988). The output of the model is a spatial activity profile that, ideally, resembles a human's brightness or lightness distribution in response to the corresponding stimulus. Several researchers have criticized the filling-in framework for assuming a "look alike" linking hypothesis (Teller, 1980) that is not logically necessary (Dennett, 1991; Kingdom & Moulden, 1989; O'Regan, 1992; Ratliff & Sirovich, 1978). Pessoa, Thompson, and Noë (1998) argue, however, that filling-in need not imply Cartesian materialism, or a *homunculus*, and thus should *not* be viewed as an isomorphism producing internal "images." Moreover, these authors stress that the debate needs to concentrate on the experimental evidence, or lack thereof, for spatially organized forms of representation. For related psychophysical evidence, see De Valois, Webster, and De Valois (1986), Hahn and Paradiso (1995), Paradiso and Nakayama (1991), and Rossi and Paradiso (1996).

Previous filling-in models have been criticized for being based on the "misguided assumption that after the brain has arrived at a discrimination or judgment, it re-presents the material on which its judgment is based, for the enjoyment of an audience in the Cartesian Theater [a homunculus], filling in the colors" (Dennett, 1991, p. 344). The type of mechanisms being advocated here are such that contrast measures are not direct correlates of perceived lightness (as in the Grossberg & Todorović, 1988, proposal), but are such that boundaries, and their associated contrast measures, provide *relational constraints* that are satisfied globally. Hence, integration is the mechanism by which color judgments are made, and not a gratuitous re-presentation of a previously made judgment. Integration, with the associated perceptual filling-in, is essential for the communication of contrast signals both within a a bounded region and across regions. It is this communication of information that allows local contrast relations to drive the system into a global lightness solution. These points are further elaborated elsewhere (Pessoa & Neumann, 1998; Pessoa et al., 1998; Ross, 1998).

In systems involving feedback or iterative processes, the speed of convergence is an important issue. This is especially relevant for vision problems, since neurons are slow processing elements, relative to electronic elements, and yet human vision occurs rapidly. Our model offers a significant synthesis of multiscale and lateral contrast integration approaches within a single algorithm that may allow rapid convergence to equilibrium, using a plausible range of retinal receptive field sizes, and still explain temporal data suggesting lateral spreading of information or filling-in. However, the plausibility of our integration scheme remains to be verified by more data on the neurophysiology of cortical multiscale, lateral, and feedback interactions.

White's Effect: Assimilation or Contrast?

From the standpoint of classical lateral inhibition proposals, White's effect presents a striking violation of simultaneous contrast as the contour length of the gray patches is larger for the stripes they do not lie on. One way to account for the effect is in terms of the directional properties of the grating (Figure 16). For a given gray patch on a black or a white stripe in the stimulus, there are two main regions of interest that determine its appearance. One is characterized by the longer contour parallel to the grating. The other by the shorter contour orthogonal to it. If for some reason, the parallel and orthogonal regions have differential effects on the appearance of the gray patch, White's illusion can occur. For example, White (1979) himself proposed that a grating might have the effect of enhancing assimilation (or reducing contrast) across borders parallel to it and reducing assimilation (or enhancing contrast) across borders orthog-

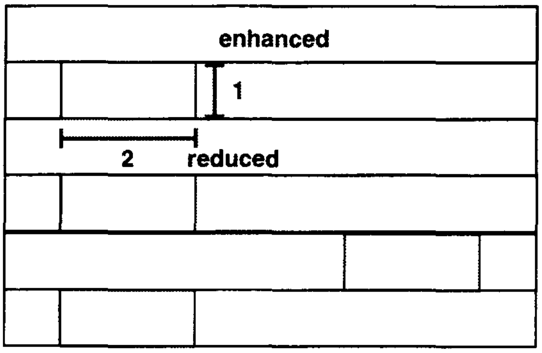

Figure 16. Stimulus for White's effect. The lightness of a given gray patch is determined by two main regions—namely, a region oriented parallel to the grating (whose common contour with the patch is indicated by 1 and a region orthogonal to the grating (contour indicated by 2). If the lateral inhibition contributions of 1 and 2 can be somehow enhanced and reduced, respectively, White's effect can occur.

onal to it. He later (White, 1981) proposed a scheme of *pattern-specific inhibition* based on elongated cortical receptive fields as a potential substrate of these ideas.

Moulden and Kingdom (1989; Kingdom & Moulden, 1991) have proposed a similar scheme but have emphasized that simultaneous contrast mechanisms are the main factor producing the effect, not assimilation. They propose a dual-mechanism model comprising a spatially local and a spatially extensive contrast effect. The spatially local effect operates all along the borders of the test patch but produces a particularly strong signal in the corner intersections of the test patch with the parallel, or coaxial bar. The more spatially extensive mechanism operates to allow the entire length of the coaxial bar to exert an influence on the lightness of the test patch. Both mechanisms operate to give the coaxial bars a disproportionate weighting in their contrasting effect on the gray test patches. In this way, White's illusion ensues. The first mechanism is thought to be subserved by center–surround receptive fields, such as those of retinal ganglion cells. The second mechanism suggests the involvement of receptive fields with small center regions and elongated surrounds.

Support for the role of *depth* in White's effect comes from the variation of White's illusion presented by Spehar et al. (1995; see Figure 2). These authors note that a display producing White's illusion can be changed to exhibit simultaneous contrast through a binocular manipulation in which gray patches are seen either at the same depth plane as a white background or at the same depth plane as a black background.

A recent study has quantified the effect of depth on White's illusion (Taya, Ehrenstein, & Cavonius, 1995). The main result of this study was that a depth organization (through binocular disparity) of a standard display of White's illusion that is consistent with our segmentation interpretation not only preserves the effect but can intensify it. According to Taya et al., the use of stereoscopic depth manipulates how display regions are grouped so that "the lightness of this rectangle [gray patch] is determined mainly by the luminance ratio between the black background and the rectangle, because they appear to lie in the same plane" (p. 692). In order to account for the Taya et al. data, we need to assume that the strength of the context boundaries produced by the stereoscopic manipulation is enhanced relative to the ones of the flat display. This way, the contributions of the coaxial bars will have an even greater weight on the determination of the lightness of the patches, when compared with the flat display situation.

Lightness Anchoring Effects

Studies such as that of the staircase Gelb effect have prompted Gilchrist and colleagues (Gilchrist et al., 2000) to advance the idea that the chief process determining lightness perception is anchoring. More specifically the perceived lightness of a given surface is given by a weighted average of the lightness of the surface when anchored relative to its local framework and the lightness of the surface when anchored relative to the global framework, where a framework is a group of surfaces that belong together. Thus, grouping plays a central role in Gilchrist's new theory, and it is proposed that factors such as coplanarity, classical Gestalt concepts such as proximity, good continuation, common fate, and similarity, as well as T-junction information, are among the determinants of scenic groupings. Within such a paradigm, they have proposed explanations of several effects, including several dealt with in the present paper, as for instance, the folded-card Adelson stimulus.

The model proposed in the present paper shares with the Gilchrist proposal a number of features—most important, the use of scenic segmentations that differently affect the lightness of a given region. Our scheme, however, differs in several important respects from Gilchrist's theory. SIM is based on the idea that filling-in instantiates a selective lightness integration process. In this context, *selective lightness integration* is the chief concept of the model. Although the model can account for some of the data that inspired Gilchrist and colleagues to defend anchoring, it does so not by explicit *anchoring computations*, but by how it selectively integrates lightness signals. Future experiments should be able to clarify whether the central process of lightness perception is anchoring (as espoused by Gilchrist and colleagues), lightness integration (as defended by the present authors), or a concept not yet developed theoretically.

Transparency and Illumination Perception

Gilchrist et al. (1983) proposed that edges are classified as illumination or reflectance edges and that 3-D surface layout must be determined before ratios are used to finally specify lightness. We propose that the separation of the scene into contexts by such cues as T-junctions affords selective integration, thereby partially discounting the effects of illumination differences. The effects of shadows or transparent films are treated together as multiplicative influences on lightness that can be partially negated through the suppression of the contrast signals at the borders between illumination contexts.[3]

The present approach does not totally discount the effects of illumination. As has been advanced by Barrow and Tanenbaum (1978, 1986) and recently advocated by Arend (1994), a full model of illumination and transparent surface perception would need to allow multiple overlapping representations—including illumination, reflectance, and transparency *layers.* At the moment, it is not clear how the visual system accomplishes this multilayer lightness representation. The current one-layer version of our model proved satisfactory for modeling the percepts we address herein. We believe it provides a working framework for the development of more sophisticated schemes that allow the characterization of illumination perception, as well as shading and transparency.

The interplay of illumination and lightness perception has gained recent attention owing to Adelson's (1993) demonstrations (see Figures 10 and 11). Adelson's original explanation of his folded-card display was related to the perception of different levels of illumination falling on the different planes. Thus, the top patch is seen as a dark gray patch that is brightly lit, whereas the bottom patch is seen as a light gray patch that is dimly lit. In contrast, in the rotated folded card, the two patches are perceived as lying in the same plane with the same illumination.

The present account of Adelson's stimulus does not invoke the perception of illumination and 3-D structure directly. Instead, we suggest that the contextual, or segmentation, factors related to the geometrical arrangement (T-junctions) of the display are responsible for the effect. This view is in line with the proposal by Todorović (1997), who was also able to devise two related displays that illustrate counterexamples to the type of explanation advanced by Adelson. We do not wish to imply, however, that illumination and 3-D arrangement factors play no role in lightness perception in general. Quite the contrary, these have been demonstrated in a number of studies involving 3-D manipulations (e.g., Gilchrist, 1977, 1980; Hochberg & Beck, 1954) as well as stereoscopic (Schirillo, Reeves, & Arend, 1990) and simulated 3-D shape effects (Pessoa et al., 1998).

Segmentation and Lightness

The chief proposal of the present work is that lightness perception is a highly context-sensitive process whose working is intimately related to image segmentation. We are, of course, not the first to advance such notions, and our approach shares several important aspects with both classic Gestalt accounts of lightness (Koffka, 1935) and more recent proposals (Gilchrist, 1977, 1980; Gilchrist et al., 2000). To illustrate our model, we have implemented a simple context–boundary module. Image segmentation is produced by a T-junction-guided process that generates image contexts that determine how lightness is selectively integrated. As has been stated, such computations were not suggested as a solution to the *image segmentation problem.* Image segmentation, in general, certainly depends on a host of cues, including motion, depth, and texture—monocular as well as binocular. The T-junction-guided scheme was chosen as an example of a mechanism that is rich enough to illustrate the ties between segmentation and lightness.

The model claims that image segmentation modulates lightness integration. But the precise source of the segmentation is not critical for the working of the model. For instance, disparity-based algorithms could have been incorporated into the context–boundary module to enable it to handle the binocular displays of Figures 2 and 3. We opted *not* to incorporate other segmentation computations, in order to concentrate on what we believe is the model's central contributions—namely, exploring a mechanism of lightness integration that is selectively controlled and showing how segmentation and lightness are tied together.

The assumption that image segmentation is provided externally for a number of our simulations expands the explanatory power of the model without having to solve, or explicitly address, the outstanding issue of how to segment images. As more comprehensive context–boundary modules are developed—for instance, incorporating disparity cues—they can be, one hopes, incorporated into the model. They simply replace assumed boundary signals with internally computed ones. In such a scenario, the context–boundary module is changed, but the bulk of the model remains. What is more, the parts that remain constitute the central elements of the model.

Shortcomings of the Model

Admittedly, the present model exhibits a number of shortcomings. For one, it proposes that lightness depends on image segmentation but only employs a primitive T-junction-based segmentation scheme. As has been stated, this step was motivated in order to concentrate the modeling effort on the *interaction* between lightness and scenic contexts. We wanted to keep the model simple. Nevertheless, the model *assumes* that, in general, context boundaries can be identified (by some as yet unspecified computations). This is certainly a strong assumption and constitutes an important element in evaluating the model.

The formalization included in this paper cannot handle dynamic aspects of brightness/lightness perception. However, in the past, filling-in was shown to be able to account for some of these data (Arrington, 1994), and we are confident that the present model will also be able to handle temporal data.

More challenging are a number of important lightness phenomena that the model has not yet been applied to. Double increment stimuli, as well as double decrement stimuli (Spehar et al., 1995) and Whishart, Frisby, and Buckley's (1995) variation on the strength of Adelson's folded-card effect as a function of the perceived angle of the folds, all remain to be simulated. Also, for the Agostini and Proffitt (1993) motion displays (see Figure 12), the model correctly assigns the direction of the lightnesses of the moving gray patches but errs in generating two background lightness values, an effect not reported in Agostini and Proffitt's studies.[4] It must be pointed out also that the magnitude of the effects generated by SIM do not always match those obtained experimentally. For instance, the magnitude of the haploscopic contrast effect produced by SIM is much smaller than that obtained by Whittle and Challands (1969). Also, the simulation of Gilchrist's coplanarity study (Figure 13) produced a much smaller effect. We contend, however, that at this stage of the development of lightness models and theories, the type of qualitative matches generated by SIM constitutes an important step.

CONCLUSIONS

This paper presents the *selective integration* neural network *model* (SIM) of lightness perception as an attempt to account for a range of lightness stimuli that have proved difficult to explain by previous accounts, such as the Benary cross, White's illusion, and Adelson's folded-card and transparency stimuli, as well as motion, depth, haploscopic, and Gelb induced contrast effects, among others. SIM offers a new treatment of how local measures of luminance contrast can be selectively integrated by a filling-in process to generate lightness percepts. Within SIM, integration between *contexts* is selectively reduced. All image boundaries do not signal the same types of scenic events. Although some outline regions of uniform coloration on surfaces, others define the edges or borders between objects, surfaces, or illumination contexts. Selectively reducing integration across these borders helps prevent the mixing of information across contexts. Selective integration can thus act to reduce the impact of illumination context differences, such as those introduced by shadows, spotlights, shading, and transparency, as well as depth changes, on perceived lightness. Furthermore, selective integration emphasizes the stable luminance relationships within objects over those variable relationships between an object and changing backgrounds.

REFERENCES

Adelson, E. (1993). Perceptual organization and the judgement of brightness. *Science*, **262**, 2042-2044.

Agostini, T., & Proffitt, D. (1993). Perceptual organization evokes simultaneous lightness contrast. *Perception*, **22**, 263-272.

Anderson, B. L. (1997). A theory of illusory lightness and transparency in monocular and binocular images: The role of contour junctions. *Perception*, **26**, 419-453.

Arend, L. E. (1973). Spatial differential and integral operations in human vision: Implication of stabilized retinal image fading. *Psychological Review*, **80**, 374-395.

Arend, L. E. (1994). Surface colors, illumination, and surface geometry: Intrinsic-image models of human color perception. In A. L. Gilchrist (Ed.), *Lightness, brightness, and transparency* (pp. 159-213). Hillsdale, NJ: Erlbaum.

Arend, L. E., Buehler, J. N., & Lockhead, G. R. (1971). Difference information in brightness perception. *Perception & Psychophysics*, **9**, 367-370.

Arend, L. E., & Goldstein, R. (1987). Simultaneous constancy, lightness, and brightness. *Journal of the Optical Society of America A*, **4**, 2281-2285.

Arend, L. E., & Spehar, B. (1993). Lightness, brightness, and brightness contrast: 1. Illuminance variation. *Perception & Psychophysics*, **54**, 446-456.

Arrington, K. (1994). The temporal dynamics of brightness filling-in. *Vision Research*, **34**, 3371-3387.

Barrow, H. G., & Tanenbaum, J. (1978). Recovering intrinsic scene characteristics from images. In A. R. Hanson & E. M. Riseman (Eds.), *Computer vision systems* (pp. 3-26). New York: Academic Press.

Barrow, H. G., & Tanenbaum, J. (1986). Computational approaches to vision. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 38.1-38.70). New York: Wiley.

Benary, W. (1924). Beobachtungen zu einem Experiment über Helligkeitskonstrast. *Psychologische Forschung*, **5**, 131-142. [Translated as "The influence of form and brightness contrast," in W. Ellis (Ed.), *A source book of Gestalt psychology*, 1939, London: Routledge and Kegan Paul].

Blake, A. (1985). Boundary conditions for lightness computation in Mondrian world. *Computer Vision, Graphics, & Image Processing*, **32**, 314-327.

Burt, P., & Adelson, E. H. (1983). The Laplacian pyramid as a compact image code. *IEEE Transactions on Communication*, **COM-31**, 532-540.

Cataliotti, J., & Gilchrist, A. (1995). Local and global processes in surface lightness perception. *Perception & Psychophysics*, **57**, 125-135.

Cohen, M. A., & Grossberg, S. (1984). Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance. *Perception & Psychophysics*, **36**, 428-456.

Dennett, D. (1991). *Consciousness explained.* Boston: Little, Brown.

De Valois, R., Webster, M., & De Valois, K. (1986). Temporal properties of brightness and color induction. *Vision Research*, **26**, 887-897.

Enroth-Cugell, C., & Robson, J. G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology*, **187**, 517-522.

Fiorentini, A., Baumgartner, G., Magnussen, S., Schiller, P., & Thomas, J. (1990). The perception of brightness and darkness: Relations to neuronal receptive fields. In L. Spillmann & J. Werner (Eds.), *Visual perception: The neurophysiological foundations* (pp. 129-161). San Diego, CA: Academic Press.

Furman, G. G. (1965). Comparison of for subtractive and shunting lateral-inhibition in receptor-neuron fields. *Kybernetik*, **2**, 257-274.

Gerrits, H., & Vendrik, A. (1970). Simultaneous contrast, filling-in process and information processing in man's visual system. *Experimental Brain Research*, **11**, 411-430.

Gilchrist, A. L. (1977). Perceived lightness depends on perceived spatial arrangement. *Science*, **195**, 185-187.

Gilchrist, A. L. (1980). When does perceived lightness depend on perceived spatial arrangement? *Perception & Psychophysics*, **28**, 527-538.

Gilchrist, A. L. (1988). Lightness contrast and failures of constancy: A common explanation. *Perception & Psychophysics*, **43**, 415-424.

Gilchrist, A. L. (1994). Introduction: Absolute versus relative theories of lightness perception. In A. L. Gilchrist (Ed.), *Lightness, brightness, and transparency* (pp. 1-34). Hillsdale, NJ: Erlbaum.

Gilchrist, A. [L.], Delman, S., & Jacobsen, A. (1983). The classification and integration of edges as critical to the perception of reflectance and illumination. *Perception & Psychophysics*, **33**, 425-436.

Gilchrist, A. L., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., Spehar, B., & Szura, J. (2000). *A new theory of lightness perception.* Manuscript in preparation.

Gogel, W. C., & Mershon, D. H. (1969). Depth adjacency in simultaneous contrast. *Perception & Psychophysics*, **5**, 13-17.

Grossberg, S. (1970). Neural pattern discrimination. *Journal of Theoretical Biology*, **27**, 291-337.

Grossberg, S., & Mingolla, E. (1985a). Neural dynamics of form per-

ception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, **92**, 173-211.
Grossberg, S., & Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception & Psychophysics*, **38**, 141-171.
Grossberg, S., Mingolla, E., & Ross, W. D. (1997). Visual brain and visual perception: How does the cortex do perceptual grouping? *Trends in Neurosciences*, **20**, 106-111.
Grossberg, S., Mingolla, E., & Williamson, J. (1995). Processing of synthetic aperture radar images by a multiscale boundary contour system and feature contour system. *Neural Networks*, **8**, 1005-1028.
Grossberg, S., & Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics*, **43**, 241-277.
Hahn, S., & Paradiso, M. (1995). Evidence for a cortical filling-in process [Abstract]. *Investigative Ophthalmology & Visual Science*, **36**(4, Suppl.), S471.
Hering, E. (1964). *Outlines of a theory of the light sense* (L. Hurvich & D. Jameson, Trans.). Cambridge, MA: Harvard University Press. (Original work published 1874)
Hochberg, J., & Beck, J. (1954). Apparent spatial arrangement and perceived brightness. *Journal of Experimental Psychology*, **47**, 263-266.
Hodgkin, A. L. (1964). *The conduction of the nervous impulse*. Liverpool: Liverpool University Press.
Horn, B. K. P. (1974). Determining lightness from an image. *Computer Graphics & Image Processing*, **3**, 277-299.
Jacobsen, A., & Gilchrist, A. (1988). The ratio principle holds over a million-to-one range of illumination. *Perception & Psychophysics*, **43**, 1-6.
Kingdom, F., & Moulden, B. (1989). Border effects on brightness: A review of findings, models and issues. *Spatial Vision*, **3**, 225-262.
Kingdom, F., & Moulden, B. (1991). White's effect and assimilation. *Vision Research*, **31**, 151-159.
Koffka, K. (1935). *Principles of Gestalt psychology*. London: Kegan Paul, Trench.
Land, E., & McCann, J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America*, **61**, 1-11.
Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: Freeman.
Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London: Series B*, **207**, 187-217.
Metelli, F. (1974). The perception of transparency. *Scientific American*, **230**, 90-98.
Moulden, B., & Kingdom, F. (1989). White's effect: A dual mechanism. *Vision Research*, **29**, 1245-1259.
O'Regan, J. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, **46**, 461-488.
Paradiso, M., & Nakayama, K. (1991). Brightness perception and filling-in. *Vision Research*, **31**, 1221-1236.
Pessoa, L., Mingolla, E., & Arend, L. E. (1996). The perception of lightness in 3-D curved objects. *Perception & Psychophysics*, **58**, 1293-1305.
Pessoa, L., & Neumann, H. (1998). Why does the brain fill-in? *Trends in Cognitive Sciences*, **2**, 422-424.
Pessoa, L., & Ross, W. (1995). A neural network model of 3-D lightness perception. (Tech. Rep. No. CAS/CNS-TR-95-016). Boston: Boston University.
Pessoa, L., & Ross, W. (1996). A contrast/filling model of 3D lightness perception. In D. Touretzky, M. Mozer, & M. Hasselmo (Eds.), *Advances in neural information processing systems 8* (pp. 844-850). Cambridge, MA: MIT PRess.
Pessoa, L., Thompson, E., & Noë, A. (1998). Finding out about filling-in: A guide to perceptual completion for visual science and the philosophy of perception. *Behavioral & Brain Sciences*, **21**, 723-802.
Ratliff, F., & Sirovich, L. (1978). Equivalence classes of visual stimuli. *Vision Research*, **18**, 845-851.
Reid, R., & Shapley, R. (1988). Brightness induction by local contrast and the spatial dependence of assimilation. *Vision Research*, **28**, 115-132.
Rock, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.
Rodieck, R. W. (1965). Quantitative analysis of cat retinal ganglion cell responses to visual stimuli. *Vision Research*, **5**, 583-601.
Ross, W. D. (1998). Filling-in while finding out: Guiding behavior by representing information. *Behavioral & Brain Sciences*, **21**, 770.
Rossi, A., & Paradiso, M. (1996). Temporal limits of brightness induction and mechanisms of brightness perception. *Vision Research*, **36**, 1391-1398.
Schirillo, J. [A.], Reeves, A., & Arend, L. [E.] (1990). Perceived lightness, but not brightness, of achromatic surfaces depends on perceived depth information. *Perception & Psychophysics*, **48**, 82-90.
Schirillo, J. A., & Shevell, S. K. (1993). Brightness contrast from a complex surround requires a complex description [Abstract]. *Investigative Ophthalmology & Visual Science*, **34**, 219.
Shapley, R., & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain controls. In N. Osborne & G. Chader (Eds.), *Progress in retinal research* (pp. 263-346). Oxford: Pergamon Press.
Shapley. R., & Reid, R. (1985). Contrast and assimilation in the perception of brightness. *Proceedings of the National Academy of Sciences*, **82**, 5983-5986.
Spehar, B., Gilchrist, A., & Arend, L. (1995). The critical role of relative luminance relations in White's effect and grating induction. *Vision Research*, **35**, 2603-2615.
Sperling, G. (1970). Model of visual adaptation and contrast detection. *Perception & Psychophysics*, **8**, 143-157.
Sperling, G., & Sondhi, M. M. (1968). Model for visual luminance discrimination and flicker detection. *Journal of the Optical Society of America*, **58**, 1133-1145.
Taya, R., Ehrenstein, W. H., & Cavonius, C. R. (1995). Varying the strength of the Munker–White effect by stereoscopic viewing. *Perception*, **24**, 685-694.
Teller, D. Y. (1980). Locus questions in visual science. In C. Harris (Ed.), *Visual coding and adaptability*. Hillsdale, NJ: Erlbaum.
Todorović, D. (1997). Lightness and junctions. *Perception*, **26**, 379-394.
Wallach, H. (1948). Brightness constancy and the nature of achromatic colors. *Journal of Experimental Psychology*, **38**, 310-324.
Walraven, J., Enroth-Cugell, C., Hood, D. C., MacLeod, D. I. A., & Schnapf, J. L. (1990). The control of visual sensitivity: Receptoral and postreceptoral processes. In L. Spillman & J. Werner (Eds.), *The neurophysiological foundations of visual perception*. San Diego: Academic Press.
Whishart, K., Frisby, J., & Buckley, D. (1995). The role of contour in a lightness illusion [Abstract]. *Investigative Ophthalmology & Visual Science*, **36**, S640.
White, M. (1979). A new effect of pattern on perceived lightness. *Perception*, **8**, 413-416.
White, M. (1981). The effect of the nature of the surround on the perceived lightness of grey bars within square-wave test gratings. *Perception*, **10**, 215-230.
Whittle, P. (1992). Brightness, discriminability, and the "Crispening Effect." *Vision Research*, **32**, 1493-1507.
Whittle, P. (1994). The psychophysics of contrast brightness. In A. L. Gilchrist (Ed.), *Lightness, brightness, and transparency* (pp. 35-110). Hillsdale, NJ: Erlbaum.
Whittle, P., & Challands, P. (1969). The effect of background luminance on the brightness of flashes. *Vision Research*, **9**, 1095-1110.
Wist, E. R. (1974). Mach bands and depth adjacency. *Bulletin of the Psychonomic Society*, **3**, 97-99.
Wist, E. R., & Susen, P. (1973). Evidence for the role of post-retinal processes in simultaneous contrast. *Psychologische Forschung*, **36**, 1-12.
Wolff, W. (1933). Über die kontrasterregende Wirkung der tranformierten Farben. *Psychologische Forschung*, **18**, 90-97.
Yarbus, A. L. (1963). *Eye movements and vision*. New York: Plenum.

NOTES

1. Schirillo and Shevell (1993) discuss another environmental challenge to constancy—namely, change in object shape. This was called Type III constancy by Pessoa, Mingolla, and Arend (1996), who evaluated its extent with side-illuminated computer-generated three-dimensional

(3-D) ellipsoid shapes. We suggest the term shape-independent constancy for this type of constancy.

2. In practice, model simulations were greatly accelerated by downsampling the lower resolution information in X^2 and the field I before computing SI. In all the simulations shown in this paper, we used 25 iterations of the large scale downsampled to ⅓ image size. For this number of iterations, lightness values closely approached equilibrium but some degree of local contrast effects were preserved. It would be interesting to further analyze whether pre-equilibrium integration results could be used as a viable means of modeling simple structural effects on perceived contrast, such as those of shallow background gradients.

3. It should be pointed out that transparency, as studied by Metelli (1974) and others, exhibits both a multiplicative and an additive, or veil, component, calling for, in general, more sophisticated processes to account for it.

4. We claim that this is due to the difference in the types of displays used. In the case of Agostini and Proffitt's (1993) displays, the background effect would presumably be there, but in a much reduced form, given the large collection of dots.

(Manuscript received November 15, 1996;
revision accepted for publication March 22, 1999.)