# A Theory of Illusory Lightness and Transparency in Monocular and Binocular Images: The Role of Contour Junctions

Barton L. Anderson
*Massachusetts Institute of Technology, Cambridge, MA  02139*

### Abstract

A theory of illusory transparency and lightness is described for monocular and binocular images containing X-, T-, and I- contour junctions. This theory asserts that the geometric and luminance relationships of contour junctions induce illusory transparency and lightness percepts by causing a phenomenal scission of a homogenous luminance into multiple contributions. Specifically, it is argued that a discontinuous change in contrast along aligned contours that preserve contrast polarity induces a scission of the lower contrast region into a near transparent surface or an illumination change, and a more distant surface that continues behind this near layer. This scission is assumed to cause changes in perceived lightness and/or surface opacity. Discontinuous changes in contrast along contours also are assumed to induce end-cut illusory contours that run roughly perpendicular to the inducing contour's orientation, both monocularly and binocularly. Binocular illusory contours are shown to be caused by the presence of unmatchable contour terminators. It is argued that the theory presented here can provide a unified account of a variety of monocular and binocular illusions that induce uniform transformations in perceived lightness, including neon-color spreading, the Munker-White illusion, Benary's illusion, and illusory monocular and binocular transparency.

## Introduction

One of the most challenging problems facing theories of perceptual organization is in determining the surface properties that generate image structure. Here, the term"surface properties" refers to the underlying causes of the image formation process, such as illumination conditions, surface reflectance, opacity, and shape. Recovering these distinct surface properties is difficult because the mapping of the 3D scenes onto 2D images is many-to-one, which implies that there is no simple means of inverting the image formation process. Yet we rarely have phenomenal access to this ambiguity; the properties of the world usually seem unambiguous and stable. This apparent stability in the face of noninvertability suggests that the visual system imposes heuristics, rules, or constraints to recover surface properties. If this logic is correct, then a fundamental problem in understanding perceptual processing is to determine what these heuristics are, and how they shape our perceptual experience of surface structure.

In this paper, I will develop a theory of perceptual organization of a family of images that contain abrupt luminance discontinuities (contours) and contour junctions. One of the fundamental ambiguities created by luminance discontinuities is in determining their distal causes. Luminance discontinuities can be generated in a variety of ways, including the occluding boundaries of objects; or by abrupt illumination changes, such as those along shadow boundaries. Despite this ambiguity, there may be systematic image properties that could provide a unique "signature" as to the environmental cause of a discontinuity. Consider, e.g., the way that contrast polarity behaves along edges. The contrast relationships generated by a shadow boundary preserves contrast polarity (since it is always darker within the shadow; cf. Cavanagh & Leclerc, 1989).[1] In contradistinction, occluding edges often generate contrast reversals, since a partially occluded background may contain regions that are both lighter and darker than the occluding surface. Similarly, transparent surfaces and changes in the optical medium -- such as fog, mist, or smoke -- will typically generate consistent polarity relationships along any edges that might arise by such media. Changes in illumination, shadows, and/or in the optical medium all correspond to image properties generated by a distinct causal layer in the image formation process, i.e., they are image properties that are not generated by the underlying surface.

Therefore, in order for the visual system to recover the underlying image causes, it needs to contain mechanisms that are capable of distinguishing the different causal layers that contribute to image structure.

The preceding considerations describe what I take to be the computational goal of the visual system which provides the rationale for the existence of the rule of image decomposition that is described below. In the remainder of this paper, it will be argued that a simple rule of perceptual grouping and segmentation can implement this computational goal, and in so doing, provide a unified understanding of a variety of monocular and binocular illusions of brightness, lightness, and transparency. Specifically, I will argue that monocular and binocular versions of illusory transparency, neon color spreading, the Munker-White illusion, and Benary's illusion, all involve a scission of a region into multiple contributions (or layers). It will be argued that the geometric and photometric relationships that occur at contour junctions are the primary cause of scission. The main proposition developed in this paper may be stated as follows:

> *When two aligned contours undergo a discontinuous change in the magnitude of contrast, but preserve contrast polarity, the lower contrast regions is decomposed into two causal layers.*

It will be argued that the decomposition of an image region into multiple layers causes monocular and binocular lightness illusions and/or various forms of illusory transparency.

## Scope of the Theory and Definition of Terms

The proposition described above requires a that number of terms be clearly defined. Below, I consider each term in the order that they appear in the proposal:

- *aligned* refers to contours that form differentiable curves (i.e., curves that are smooth at the highest level of resolution available to the visual system). There remains a critical issue of exactly how to define *aligned*, which has been discussed previously by a number of authors in the context of contour interpolation (cf. Anderson & Julesz, 1995; Field, Hayes, & Hess, 1993; Grossberg & Mingolla, 1985; Grossberg, 1993; Kellman & Shipley, 1991). In general, we have observed that all of the effects described herein exhibit some monotonic decrease in strength as contour alignment is perturbed, which has little impact on the basic tenet of the proposal (since the proposal does not contain an explicit commitment to the quantitative dependence of scission on degree of alignment).

- a *contour* refers to a step function in a luminance profile.

- *discontinuous change in magnitude of contrast* refers to the presence of a discontinuity in a contour's derivative.

- *magnitude of contrast* refers to the size of the luminance difference across a contour. Throughout this paper, I will use the term *contrast* to refer to both chromatic and achromatic stimuli. The working assumption is that the achromatic contrast across a contour is the primary determinant of scission.[2] There also is a general issue of whether a divisive term should be included in a contrast measure (e.g., the total or mean luminance), which I will not consider in this paper. The various definitions of contrast that employ such normalization factors would apply to an entire image, and since I will consider only the relative magnitudes of contrast *within* an image, this term may be ignored.

- *contrast polarity* refers to the sign of the contrast difference across an edge. In general, this requires the imposition of a coordinate system so that the direction of contrast may be given a sign. Here, we will only consider the relative polarity relationships of aligned contour segments, such that each contour segment serves as the coordinate system that defines the polarity of the adjacent contour (either the "same" or "opposite").

- *causal layers* refers to the attribution of an image region's luminance to more than one layer (e.g., a transparent layer and an underlying surface, or the reflectance of a surface and the prevailing illumination conditions).

It is important to emphasize at the outset that the theoretical stance described here blurs the distinction between lightness transformations and relative degrees of transparency, at least in geometric contexts that do not uniquely distinguish lightness and transparency. The justification for this conceptual blurring is grounded on the assumption that observers are often incapable of distinguishing lightness and/or illumination changes from changes caused by the medium through which a surface is visible (including transparent surfaces). In the theory described below, the term *scission* is therefore used to refer to a decomposition into multiple sources, where "sources" is understood to refer to distinct causal "layers." In the present framework, the assertion that scission occurs does not require that observers have any direct *awareness* of multiple surface layers. Rather, the idea developed here is that scission may *reveal* itself as either transparency *or* a lightness transformation, and that both transformations are a consequence of relatively "dumb" mechanisms that decompose an image region into multiple causal contributions.[3]

The arguments described below should be understood to represent *sufficient* conditions for the occurrence of the illusions described within, not necessary conditions. For example, although the focus of the present paper is on images containing contours and contour junctions, there is no claim that the presence of contour junctions are *necessary* for illusory transparency or color scission to occur. Indeed, it was recently demonstrated that contours *per se* are not necessary for any of these effects (Anderson, 1995, 1996). However, we have found that the principles outlined below may be extended to handle more complex images that do not contain contours and contour junctions, and hence no loss of generality accrues from restricting attention to these (putatively) simpler image configurations. The theoretical arguments presented in this paper are also not intended to cover the entire range of brightness or lightness illusions that have been reported in the perceptual literature. Rather, I will focus on a specific set of stimulus configurations that share a number of photometric and geometric properties, so that I may precisely articulate the critical image properties that give rise to illusions of lightness or transparency in the subset of patterns containing these features. There is also no attempt to provide a detailed account of the problem of "anchoring" (see, e.g., Gilchrist, Kossyfidis, Bonato, Agostini, Cataliotti, Li, Spehar, & Szura, 1996), or the production of spatially inhomogeneous brightness inductions such as Mach bands or grating induction (McCourt, 1982; McCourt & Blakeslee, 1994). Rather, in this paper, I will restrict attention to a variety of illusions that lead to a uniform transformation of the perceived brightness or opacity of a given homogenous luminance patch.

This paper is divided into five sections. In Section 1, the concept of scission is introduced in the context of phenomenal transparency in monocular images that contain X-junctions. Sections 2 and 3 extend the principle of scission to monocular images containing T-junctions and I-junctions that generate percepts of transparency or lightness illusions. In Section 4, a binocular theory of scission and illusory contour formation is described. Section 5 compares the present theory with previous models and concludes with a discussion of the limitations of the present perspective and suggestions for future research.

## Section 1: Color Scission, Transparency, and X-junctions

One of the most ardent advocates of scission as a principle of perceptual organization was Metelli (1970, 1974a,b; Metelli, da Pos, & Cavedon, 1985). Metelli derived an extremely influential quantitative model of transparency that was based on a generalization of Talbot's law. Talbot's law outlined the conditions that must be met for color fusion, i.e., the fusion of multiple frequencies into the appearance of a single color. Metelli (1970, 1974a,b) suggested that Talbot's law could serve as a model of achromatic transparency if the equations used to derive color fusion were understood as also describing the decomposition of a single achromatic luminance into two sources. In Metelli's theory, the concept of scission was used to describe the phenomenal decomposition of a single region of uniform luminance into two surface layers, one of which was transparent. Since it was in this domain that the concept of scission was first applied, the problem of transparency serves as a natural starting point for developing the concept of scission as a general principle of perceptual organization.

Metelli's model of transparency was developed under the experimental context of an episcotister (a rotating disc with open sectors) and a bipartite background. This experimental context generates images such as those depicted in Fig. 1. Metelli's laws state that the luminance of regions p and q -- the regions containing a mixture of two surfaces -- must be the weighted average of the reflectance of the near surface (the episcotister) and the more distant surface visible behind this region. If the mixture weights are expressed as proportions rather than absolute quantities, then the regions p and q may be written:

$$p = \alpha a + (1 - \alpha)t \qquad (1)$$

$$q = \alpha b + (1 - \alpha)t \qquad (2)$$

where $p$ is the reflectance of the stimulus color, $a$ and $b$ the reflectance of the second layer, $t$ the reflectance of the transparent layer, and $\alpha$ and $(1-\alpha)$ the proportions into which the stimulus color has been divided between the two layers. In these equations, $p$, $q$, $a$, and $b$ are data terms that are given in the image, and $\alpha$ and $t$ are unknowns. With some simple algebraic manipulations, these equations may be solved, yielding:

$$\alpha = (p - q)/(a - b) \qquad (3)$$

$$t = (aq - bp)/[(a + q) - (b + p)] \qquad (4)$$

In order for these equations to make any physical sense, a number of additional constraints are required. Since $\alpha$ refers to a degree of transparency of a physical surface, this implies that $\alpha$ must be between 0 and 1, which in turn implies that the numerator and denominator in equation (3) must have the same sign. The same sign constraints hold for equation 4, since $t$ represents a reflectance, and negative values would be an absurdity.

In evaluating these algebraic constraints on transparency, it is important to keep in mind that the solutions derived in equations (3) and (4) only apply to the case of "balanced" transparency, i.e., where the transparent surface's reflectance and transmittance may both be characterized by a single number across the entire transparent surface. In other words, it must be the case that the $\alpha$'s and t's in equations (1) and (2) are the same. In the experimental context of the episcotister, this assumption is true by design. However, in general, there is no *a priori* physical requirement that a transparent medium or surface have a uniform degree of transmittance and/or reflectivity. In such cases, Metelli's theory "...makes no assertions about the possible existence of unbalanced forms of transparency (Metelli et al., 1985, p. 355)." Thus, Metelli's theory is not a general theory of transparency, but rather, is restricted to those sets of images that have only 4 (or less) luminance values.

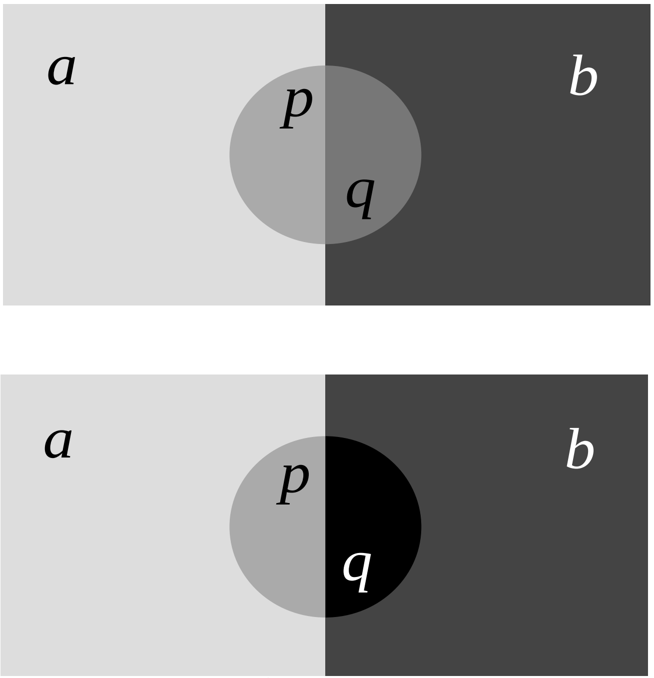


***Figure 1***

A display illustrating the quantities used in Metelli's equations. In this figure, regions $p$ and $q$ correspond to the image regions containing a mixture of two colors, and $a$ and $b$ refer to the reflectance (or lightness) values of the more distant surfaces.

Metelli's equations therefore describe some precise quantitative relationships between projected reflectance values and possible transparent surfaces in a restricted subset of displays. In principle, these quantitative relationships could be used by the visual system to recover information about the transmittance and reflectance of the transparent surface. However, given the restrictive set of physical contexts that are described by these equations, there is no reason to believe that the visual system would invoke a set of constraints that have such a narrow domain of application. After all, it is clearly possible for the transmittance and/or reflectance of a surface to be unbalanced (such as might occur with fog, clouds, or smoke). Therefore, in order to work in these more general contexts, the visual system may apply some simple rules of visual grouping and segmentation to recover transparency, rather than requiring specific quantitative relationships to hold. Such rules need not respect the quantitative relationships entailed in a physically derived model of transparency, but instead, may only reflect the qualitative boundaries of "possible" and "impossible" transparent media. One of the purposes of this paper is to develop arguments supporting the existence of just such a qualitative model.

The thesis developed here is that the perception of transparency in images that contain contour junctions is determined by the contrast relationships of the contours that are generated along the collinear segments of X-, and T-, and I-junctions (see Mackworth, 1976, for a review of the use of image junctions in scene interpretation). In particular, I will argue that contrast polarity is the primary property used by the visual system to decompose a region into multiple contributions. By definition, contrast polarity is a qualitative property that varies only in sign (positive, negative, or zero). Thus, a theory that relies on contrast polarity as a critical determinant of transparency percepts will necessarily also be qualitative in nature. Although this may be seen as a weakness of the present theory, it may also imbue this theory with a greater generality than the quantitative models of transparency perception that have been developed to date. Such quantitative models require that the visual system have access to fairly rich body of information about the physics of light and reflectance laws, including the illumination conditions, reflectances, and opacities of the surfaces in an image. Models of this kind are typically referred to as *intrinsic image models* (Barrow & Tenenbaum, 1978). To our knowledge, there has never been a principled account for how such information could be acquired by the visual system. Therefore, it may be more prudent to assume that the visual system recovers the causes of image structure by tracking a few simple image features that serve as signatures for the presence of multiple image causes, rather than require that a large amount of prior knowledge be built into the visual system for it to operate effectively.

We begin our analysis by considering the contrast relationships that arise at the contour junctions in the images studied by Metelli. Our goal is *not* to provide an exhaustive account of all patterns containing X-junctions, which would require an examination of 4! = 24 distinct luminance combinations. Rather, the goal is to analyze X-junctions so that the polarity relationships that are critical for the perception of transparency may be discerned.

The stimulus studied by Metelli generated two sets of X-junctions, an image property that has been studied previously as a cue for transparency (Adelson & Anandan, 1990; Kersten, 1991). The left side of Figure 2 portrays a variety of monocular images generating X-junctions, and the right side depicts the polarity relationships of the edges generated by the X-junctions. In order for an X-junction to be consistent with the presence of a transparent surface (or a change in illumination), the polarity of an edge underneath the transparent surface must be preserved; only the contrast magnitude can change. This follows from the simple fact that transparent surfaces (or an illumination change caused by a shadow) can only reduce the contrast of an underlying contour; the contrast polarity of edges must be preserved. In the images depicted in Fig. 2, it can be seen that the image configurations that support percepts of transparency are those that preserve the polarity of at least two of the aligned contours.

When X-junctions are classified according to the polarity relationships of aligned contours, all of the possible luminance combinations may be classified into three basic kinds of junctions. In keeping with the terminology introduced by Adelson & Anandan (1990), these junctions may be termed non-reversing, single-reversing, and double-reversing. Non-reversing X-junctions give rise to a bistable percept of transparency, wherein either or both sets of aligned contours may appear as a transparent surface (or, say, as overlapping shadows). Which surface appears transparent depends on the perceived depth order, which is also ambiguous in these monocular images (cf. Fig. 2a). Single-reversing X-junctions typically give rise to a unique assignment of transparency, as well as a unique depth ordering of the surfaces (see Fig. 2b). Double-reversing X-junctions do not give rise to percepts of transparency (and are physically impossible; see Fig. 2c).[4] If we assume that the visual system uses a combination of contrast polarity and contrast magnitude to determine the presence of a transparent surface (or a change in illumination), then all of the examples of transparency described by Metelli can be understood.

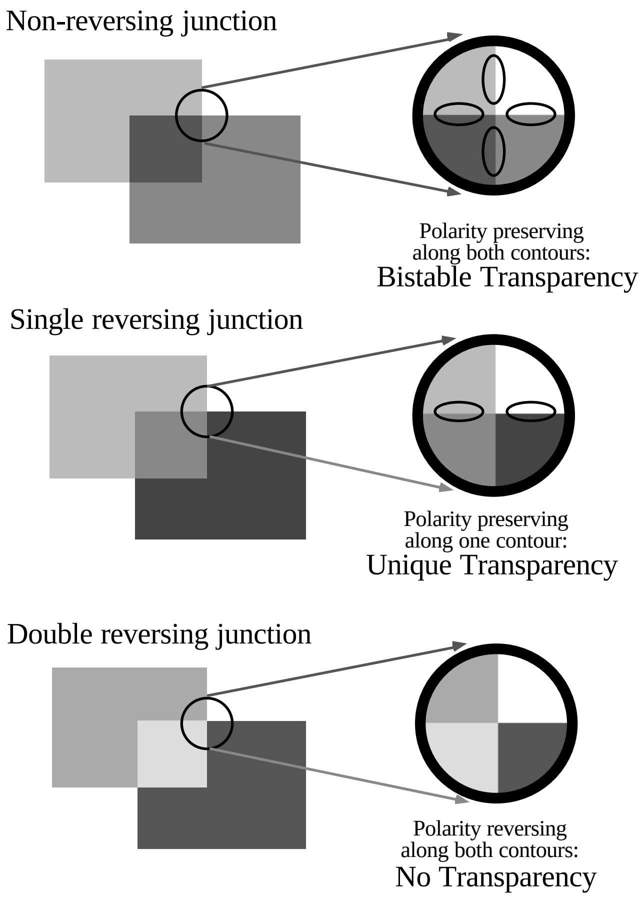


***Figure 2***

Figure depicting the kinds of X-junctions that can be generated in simple images, and the kinds of percepts these junctions support. The circular apertures are included to indicate those contours that preserve contrast polarity and those that do not. Since a transparent surface can only change the magnitude of a contour's contrast and not its polarity, only those edges that preserve contrast polarity across the X-junction are capable of supporting a transparent interpretation. In the top figure, both sets of aligned contours preserve contrast polarity. This implies that either contour may be overlaid with a transparent surface, which generates a bistability in the perceived depth of the two layers. The middle figure depicts a pattern with only a single set of contours that preserve polarity, and hence, only this contour is consistent with an overlaying transparent surface. In the bottom figure, neither pair of aligned contours preserve polarity, and hence, this pattern does not support the percept of transparency. (After Adelson & Anandan, 1990).

Thus, in general, the percept of transparency -- a phenomenal scission of a single luminance into multiple layers -- arises when at least one pair of aligned contours across an X-junction preserves contrast polarity, accompanied by a change in the magnitude of contrast. The contrast magnitude change can also uniquely specify the three-dimensional ordering of the transparent surface when just one pair of contours preserve polarity: the transparent surface must be located *over* the contour that has the smaller contrast. Again, it is important to emphasize that the same polarity preserving rule will also describe a change in illumination, such as that which occurs along shadow boundaries. Thus, the same rule predicts the occurrence of both shadow junctions and those caused by transparent surfaces, whereas occluding edges do not necessarily preserve contrast polarity.

In sum, consistent contrast polarity relationships along aligned contours can provide a compelling signature of the presence of an illumination change and/or the presence of a transparent medium. This property has been used by a number of authors to restrict the possible luminance values that can be used to reliably infer the presence of a transparent surface (cf. Adelson & Anandan, 1990; Metelli, 1974a,b; Metelli et al., 1985). The primary difference between the present analysis and those described previously is that I assume that this image property serves as a *sufficient* condition for scission. As noted above, this assumption is largely motivated by the belief that the visual system is incapable of performing a complete intrinsic image analysis, and therefore must attempt to attribute image properties to their environmental causes by applying a few principles of grouping and segmentation. In the sections that follow, I suggest that this simple rule can explain a number of illusions with a single theory, some of which have eluded explanation even to this date.

## Section 2: Color Scission in Images Containing T-junctions

In Section 1, I focused on the photometric relationships generated along X-junctions. In this section, I extend this analysis to encompass displays containing T-junctions formed by the intersection of three luminance values. There exists an extensive literature on the role of contour junctions in specifying the three-dimensional structure of images (Anderson & Julesz, 1995; Clowes, 1971; Guzman, 1968; Huffman, 1971; Mackworth, 1975; Malik, 1987). Typically, T-junctions have been treated as signaling the presence of an occlusion relationship, where the "top" of the T is assumed to occlude the "stem" of the T (see Fig. 3 for the meaning of these terms). Here, I will argue that when a T-junction is formed by three distinct luminance regions and the aligned contours along the top of the T-junction preserve contrast polarity that two processes are initiated: (1) a contour is generated that forms roughly orthogonal to the orientation of the inducing contour segments, which generates either a modal or amodal continuation of the T-junction stem; and (2) the lower contrast edge of the aligned contours is decomposed into a product of two distinct layers. It will be argued that these simple rules are capable of explaining all of the lightness illusions that are generated by images containing T-junctions that satisfy this polarity constraint.

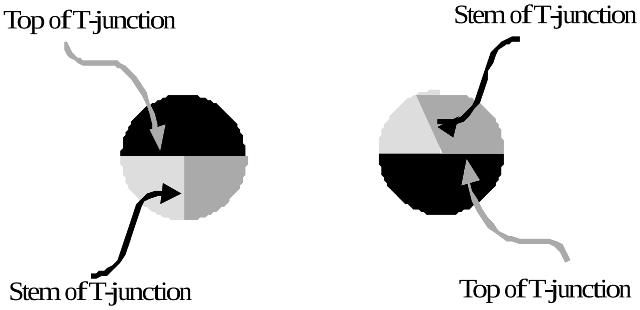


**Figure 3**

Figure illustrating the terms used to describe the components of a T-junction.

There are a number of issues that must be addressed in any theory of contour interpolation, most of which have been discussed at length previously (Anderson & Julesz, 1995; Grossberg & Mingolla, 1985; Kellman & Shipley, 1991; Field, Hayes, & Hess, 1993). In this and the following section, I simply assume that discontinuities can generate illusory contours, and do not commit to a particular mechanism of how these contours form. In part, this omission is motivated by the desire to restrict the present monograph to a reasonable length, and to emphasize the importance of scission for understanding the nature of the illusions described below. All that is needed for present purposes is the recognition that contour formation processes can be induced orthogonally to contour terminations, and that these locally generated contours may group together to form coherent shapes when certain geometric conditions are met (for a discussion of these geometric constraints, see Anderson & Julesz, 1995; Grossberg & Mingolla, 1985; Kellman & Shipley, 1991; and Field, Hayes, & Hess, 1993).[5] We will discuss how the aligned contour segments along the tops of T-junctions may be understood to terminate in Section 4 below.

Let us begin our discussion of T-junctions with a paradoxical brightness/lightness illusion[6] discovered by Munker (1970) in the chromatic domain, and subsequently by White for achromatic stimuli (1979, 1981; hereafter referred to as the Munker-White illusion; cf. Taya, Ehrenstein, & Cavonius, 1995). The Munker-White illusion is generated when grey bars are placed within more extended lighter and darker bars (see Fig. 4). This configuration causes the longer sides of the grey bars within the darker stripes to be bordered by the highest luminance, and the longer sides of the grey bars within the light stripes to be bordered by the darkest luminance. Contrast enhancement mechanisms predict that the grey bars neighboring the lighter stripes should appear darker than the grey bars neighboring the darker stripes, but the opposite effect is observed.

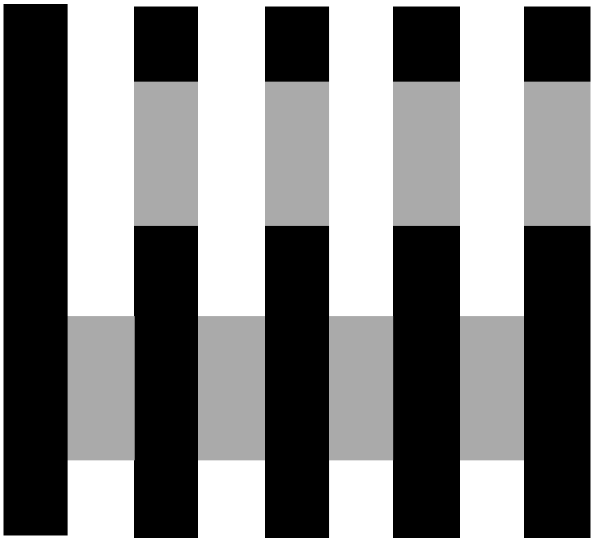
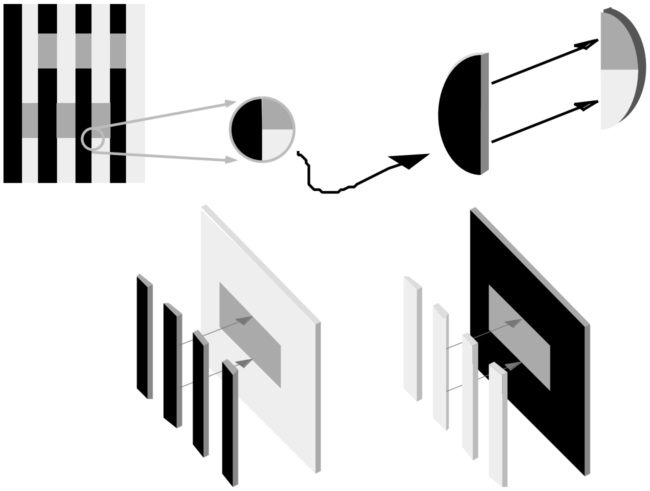


***Figure 4***

The Munker-White illusion, and the corresponding T-junctions that arise in this image.

Recently, a number of authors have suggested an account of the Munker -White illusion that builds on the interpretation of T-junctions as signaling occlusion relationships (Gilchrist et al., 1996). The basic insight of this account builds on the observation that the visual system seems to disregard the contrast generated along the longer side of the grey bars, since these contours would otherwise cause a brightness illusion opposite to that which is observed (assuming, of course, that lateral inhibition is a correct explanation of simultaneous contrast, which requires a host of assumptions to work; see, e.g., Pessoa, 1996). Therefore, a minimal requirement of any successful theory of this illusion is to provide some means for understanding how the bars neighboring the targets are discounted. The T-junction/occlusion explanation runs as follows. T-junctions usually arise from one surface occluding another. Therefore, the contour along the top of the T-junction -- the adjacent stripes -- should be interpreted as occluders, and the grey targets and the stripes in which they are embedded should be interpreted as occluded. Such a dissociation transforms the Munker-White illusion to a simultaneous contrast display (see Fig. 5), which correctly predicts the sign of the Munker-White illusion, and (apparently) restores the veracity of contrast enhancing mechanisms (although no commitment to a particular theory of simultaneous contrast is needed for this explanation to work; all that is necessary is that the two sets of displays are shown to produce the same illusion). Note, however, that this explanation of the Munker-White illusion predicts more than just the sign of the illusion. It also predicts that the Munker-White illusion should only be as *large* as simultaneous contrast, depending on the extent to which the dissociation between the grey bars and the neighboring regions is complete (see Fig. 6). If the dissociation is less than complete, then the neighboring bars should drive the illusion in the opposite direction of that which is observed. However, if the dissociation *is* complete, then this means that the Munker-White illusion has been reduced to a simultaneous contrast display, and should therefore only be as large as simultaneous contrast. But as can be seen in Fig. 6, the problem with this explanation is that *the Munker-White illusion is larger than the illusion iduced by similar simultaneous contrast displays*, which means that this explanation of the Munker-White illusion cannot be correct (or minimally, that this explanation is incomplete).



Occlusion/T-junction Theory of the
Munker-White Illusion

***Figure 5***

The T-junction/occlusion theory of the Munker-White illusion. This theory suggests that the T-junctions introduce a depth cue that places the stems of the T (the gray bars) behind the tops of the T (see Gilchrist et al., 1996). As shown on the bottom of the figure, this would effectively reduce the display to a simultaneous contrast display, and predict the correct sign of the lightness illusion.

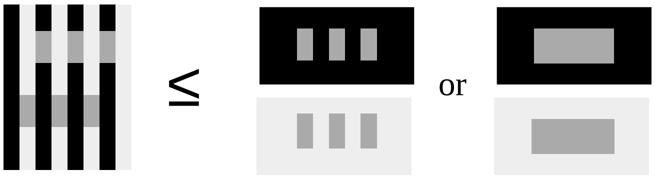


**Figure 6**

The prediction of illusion magnitude from the T-junction/occlusion analysis depicted in Fig. 5. A local interpretation of T-junctions as occluders would reduce the Munker-White display to either a set of individual grey bars a homogeneous background (middle column), or a single grey region if amodal completion occurs (far right column). This means that the magnitude of the Munker-White illusion should be less than or equal to the magnitude of one of these displays, but observers uniformly report that the Munker-White illusion is *stronger* than either of these patterns. This fact demonstrates the inadequacy of the T-junction/occlusion account.

Here, a very different explanation of this illusion is offered. Consider the luminance relationships that arise at the T-junctions in the Munker-White display. The aligned contours along the top of the T-junction preserve contrast polarity, but the contrast of one of the edges is reduced relative to that of the other. When analyzing X-junctions, we found that such changes were predictive of patterns that elicit percepts of transparency. The thesis forwarded here is that *the Munker-white illusion is the consequence of a perceptual scission that splits the lower contrast region along the top of the T into multiple sources*. For purposes of exposition, let us treat this decomposition as giving rise to a transparent surface overlaying a background, and consider the perceptual consequence of this decomposition for each set of grey bars (a similar story would hold if we treated the contrast change as due to a variation in illumination, although with a different phenomenal consequence). When the grey bars are embedded in a black stripe, the hypothesized scission mechanism will treat the grey region as a product of a continuous black stripe and a light colored filter that overlies this black stripe. But when the grey bars are embedded in a white stripe, a scission mechanism will treat the grey region as a product of a continuous white stripe and a dark colored filter that overlies this white stripe. The claim here is that this decomposition causes the grey bars in the white stripes to appear darker because some of the lightness in the grey bars is attributed to a continuation of the white stripes, rather than the grey bars themselves. In a similar vein, the grey bars in the black stripes appear lighter because some of the darkness in the grey bars is attributed to the continuation of the black stripes, rather than the grey bars themselves. This is depicted in Fig. 7.

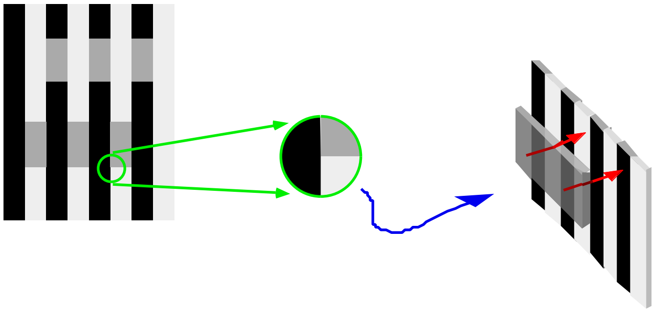


**Figure 7**

A schematic depicting the scission explanation of the Munker-white illusion. See text and Fig. 8 for details.

There is at least one tacit assumption in the preceding analysis that should be made explicit. I have argued that the aligned contours along the top of the T-junctions induces a scission of the lower contrast region into layers, but I have not committed to a particular model of how this scission occurs. In principle, there is an infinite number of ways in which this decomposition can be realized, corresponding to different attributions of lightness and/or transparency to the multiple layers. However, in the preceding discussion, I have assumed that one of the layers has the *same* lightness as the extended stripes in which the targets are embedded. This decomposition is depicted in Fig. 8. When we consider stereoscopic T-junction displays in Section 4 of this paper, we will see that it does seem that this assumption is employed by the visual system, at least when clear percepts of transparency are generated stereoscopically. A host of possible reasons can be put forth to explain *why* such an assumption would be made, but no commitment to any particular rationale is integral to the theory described in this paper.[7] However, the need for an assumption of this kind (or something computationally equivalent) is paramount for the present theory. The reason for its importance can be explained as follows. The thesis of scission presumes that a given luminance is decomposed into multiple sources. If any arbitrary decomposition is allowed, then an ambiguity emerges. Consider Fig. 8. In this example, a middle grey is decomposed into a light and a dark contribution. If observers are now asked to judge the brightness, lightness, or opacity of the target region, which layer in this region should they report? One layer would lead to a report of "lighter" and the other would lead to "darker" (relative to an undecomposed region containing a single layer). The theory would be incapable of distinguishing which direction of illusion should arise, or it might predict a bistability in illusion direction. However, if one layer is treated as a continuation of the adjacent luminances bounded by the aligned contours, then the "target" is uniquely defined: it is the layer that causes a change in the appearance of the continuous surface. This removes the ambiguity in defining the "target," yielding a specific predicted illusion direction.

Thus, it is a critical aspect of the present theory that one of the multiple layers be considered to be a continuation of the stripe in which the target is embedded, so that the thesis of scission can predict a definite sign of the illusion. Strictly speaking, this does not imply that it is necessary to assume that one of the scission colors be *identical* to the stripes in which the target is embedded. It is well known that when there is ambiguity in how to group layers within a scissioned region, that the visual system groups those contour segments that are most similar in color (Beck et al., 1984; Morinaga, Noguchi, & Osihi, 1962; Petter, 1960). Thus, all that is required for a unique prediction in the direction of the illusion is that it be recognized that one of the scission colors needs to be attributed to the continuation of the stripes in which the targets are embedded.

The scission explanation of the Munker-White effect leads to two predictions about factors that should modulate the strength of the illusion. First, scission implies that the luminance variation along the stem of the T-junction may actually diminish the magnitude of the illusion, since it creates a cue that the image regions on either side of the T-stem are behind the top of the T-junction (i.e., occluded). In contrast, the proposed scission mechanism would treat the grey region as containing a near layer, which could be interpreted as a transparent surface or an illumination change (such as a shadow). This implies that the Munker-White illusion should be strengthened by either reducing the salience of the T-stems in the images, or by providing an unambiguous cue for the presence of a near layer.

One simple method for reducing the salience of the T-junction is to reduce the width of the stripes. As shown in Fig.9, the Munker-White illusion can indeed be strengthened by this simple manipulation, as can its chromatic variant. A similar enhancement was previously noted by White (1979, 1981), but interpreted quite differently. Note that the alignment of multiple T-stems is also not necessary to produce the illusion, as shown in Fig. 9, supporting the relatively local analysis presented here.

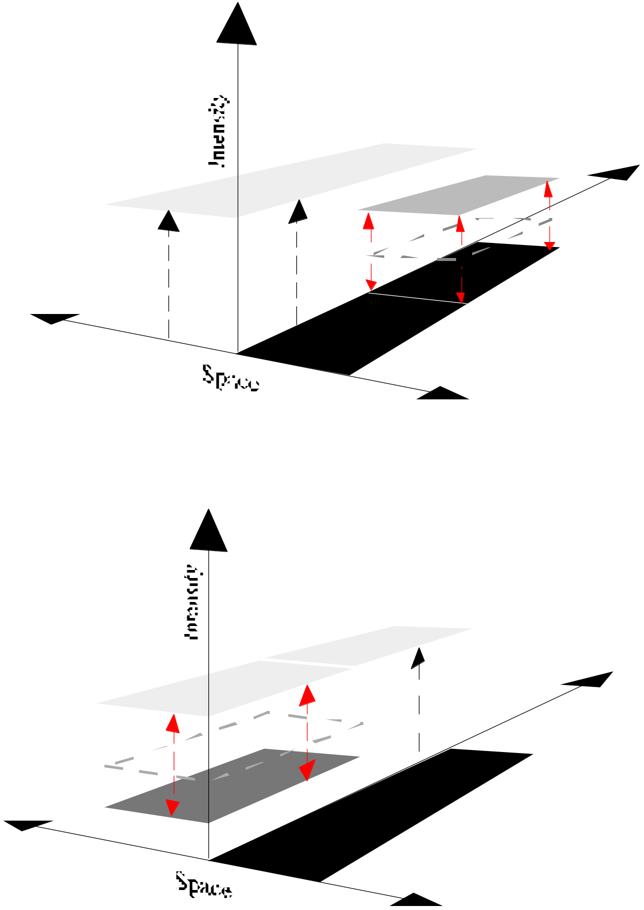


***Figure 8***

A schematic depicting the scission hypothesis as it would apply to a single target embedded within a black stripe (top) and a white stripe (bottom). In these figures, the dashed rectangular regions correspond to the actual luminance values of the grey regions. The solid grey region corresponds to the transformation in the target's color as a consequence of scission. In these figures, it is assumed that decomposition induced by scission is done is such a way that one of the layers in the scissioned region is assigned the same color as the stripe in which it is embedded.

A problem with using displays like those depicted in Fig. 9 to assess the validity of the scission account of the

Munker-White illusion is that a number of other theories would predict a similar enhancement by this manipulation. Indeed, one of the most common explanations of the Munker-White illusion is that it involves assimilation, wherein the target region is in some way "blended" or averaged with neighboring image regions. This seems like a natural explanation of this illusion, since assimilation will--by definition -- induce the opposite transformation of that produced by contrast enhancement mechanisms. The patterns presented in Fig. 9 have historically been presented as examples of spatial regimes in which assimilation has been presumed to operate, so in order to use such displays as evidence for support of the scission thesis, we must first demonstrate that an assimilation explanation fails when it makes predictions that are different from those predicted by scission.

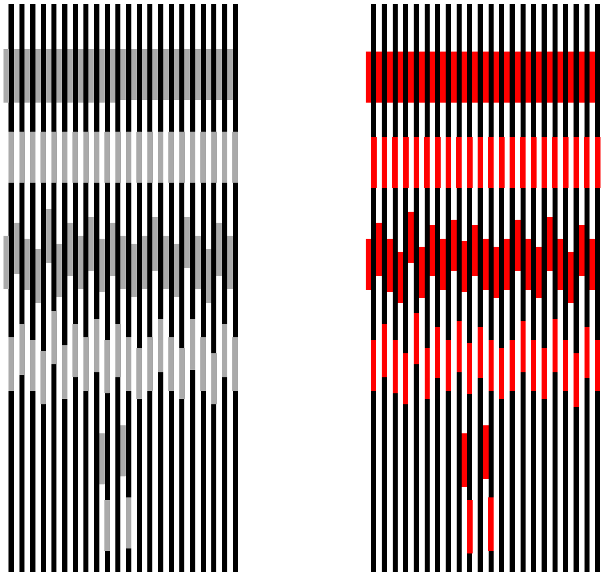


***Figure 9***

Chromatic and achromatic variants of the Munker-White illusion, in which the width of the stripes has been greatly reduced. The illusion is greatly enhanced for both columns, and do not r equire coherent subjective contours to occur, supporting the relatively local analysis described here.

This problem requires that we find a set of parameter regimes in variants of the Munker-White illusion where the assimilation explanation makes a different prediction than scission, and determine which model succeeds (if any). A series of figures that allows us to discriminate these two theories is presented in Fig. 10. The critical difference between the predictions made by assimilation and scission is that the proposed scission mechanism tracks the contrast polarity of the aligned contours, whereas assimilation is essentially "blind" to contrast polarity. More specifically, scission is only hypothesized to occur when the contrast polarity of aligned edges along the top of the T-junction have the same sign, whereas assimilation predicts that a grey bar neighboring lighter stripes should appear lighter than the same bar neighbored by darker stripes, irrespective of the polarity relationships arising along the tops of the T-junction. Fig. 10 reveals the failure of the assimilation prediction for a number of variants of White's display. In the three examples presented in this figure, assimilation predicted the opposite percept of that reported by observers: those regions predicted by assimilation to appear darker actually appeared lighter. It is also worth noting that many of the observers that viewed these patterns commented that the target regions that contain contrast reversals appeared *qualitativelydifferent* than those that preserved contrast polarity, and voiced some objections that to the idea that "lighter" and "darker" were sufficient terms to capture the differences in the target's appearance. This provides suggestive support for the contention that contrast polarity of aligned contours is a critical dimension that modulates the appearance of the targets in the Munker-White illusion.

A second prediction of the scission thesis is that the Munker-White illusion should be enhanced if the grey stripes are given a depth that causes them to appear in front of the longer light and dark bars. By hypothesis, the Munker-White illusion is the consequence of an achromatic color scission. By presenting the grey bars in front of the flanking stripes, the strength of this scission should be enhanced.[8] There have been two previous studies that have attempted to determine the role of stereoscopic depth in modulating the strength of White's illusion (Taya et all., 1995; Pessoa & Ross, 1996), but, strangely, these studies have reported some conflicting results. Both studies have reported that the Munker-White illusion is strengthened when the grey bars appear in front of the dark and light bars in the image. However, these studies differ greatly in the reported impact of placing the grey bars behind the flanking bars. One study (Taya et al.) reported that the strength of the Munker-White's illusion was increased when it was placed behind the flanking bars, albeit less than when they were placed in front. The other study (Pessoa & Ross, 1996) reported that the effect was greatly diminished (to the point of not existing) when the grey bars were placed behind the flanking bars. How can the discrepancy in these two studies be reconciled?

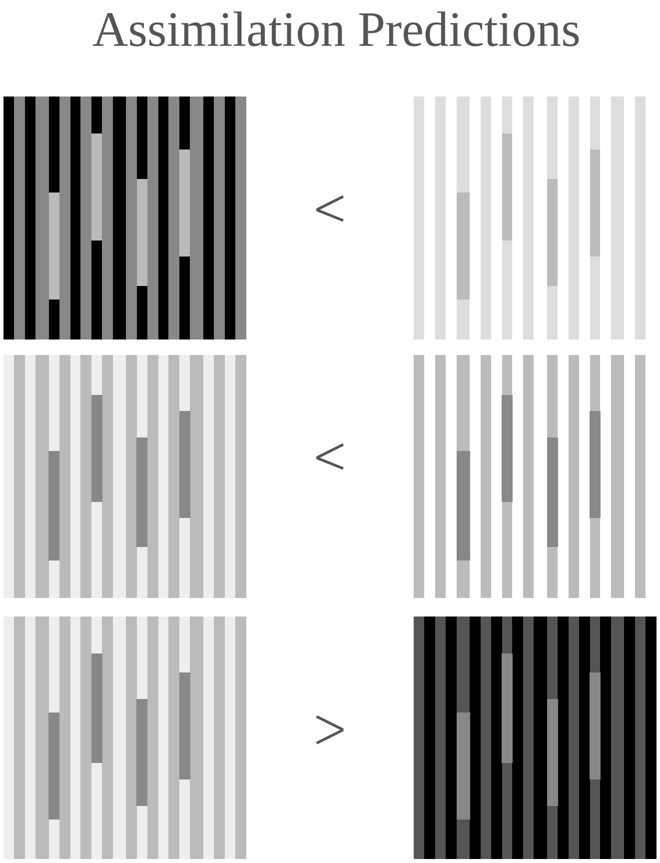


***Figure 10***

Demonstration of the failure of assimilation in correctly predicting the sign of the illusion for a number of White's patterns. The small grey bars within a row are all the same luminance. The inequalities express which of the target bars should appear lighter (where greater = lighter), under the assumption that "assimilation" means that the target bars should be a weighted average of the real luminance value and that of its neighbors. Note that in all examples depicted, that assimilation fails to predict the correct direction of perceived lightness difference. See text for details.

In order to provide some understanding of these discrepant results, we created stereo versions of these displays similar to those previously generated, but with more potent stereo cues. One problem with both previous studies is that they used rather weak forms of depth information to create the depth segregation. For example, Taya et al. used a vertical grating and introduced horizontal disparity to only the two outer bar stripes. This creates a weak form of depth information, because none of the horizontal contours (the stems of the T-junctions) generate disparity signals, and only the far left and right bars generate the appropriate disparity signal. This means that any perceived depth information attributed to the central grey bars was largely a consequence of interpolation mechanisms. We therefore reconstructed these displays using more vivid stereoscopic depth information (see Fig. 11). These images were created by displacing the grey bars along the 45 degree axis. In the Section 4, the role of such displacements in creating depth will be described; at this juncture, we only need note that such displacements are capable of eliciting strong percepts of stereoscopic depth. The grey dots in Fig. 11 were added to provide fusion "locks" to prevent observers from matching the vertical component of these shifts.

We tested seven observers and asked them to answer two questions: 1) whether the difference in the appearance of the grey bars in the near condition was larger, smaller, or equal to the difference in the appearance of the grey bars in the far condition; and 2) whether the size of the illusion in the far condition was the greater, smaller, or equal to the case in which the grey bars appeared at the same depth as the flanking bars. All of the observers that perceived stereoscopic depth in the four displays reported that the near condition produced a much stronger apparent brightness difference, with one set of grey bars appearing as a dark layer overlying light bars, and the other appearing as a light colored film overlying dark bars. Responses were mixed when observers compared the behind condition to the same plane condition. Five observers reported that the illusion was diminished or unaffected when the grey bars were behind, but the two others reported that the illusion was, in fact, enhanced.

In order to determine the mixed responses to the behind configuration, observers were asked to describe the appearance of the entire display. The observers that reported the illusion to be diminished or unchanged from the monocular version of the illusion reported that all of the luminance within the stripes in which the grey bars were embedded appeared behind the flanking stripes at the same depth as the grey bars. For example, when the grey elements were placed within the black stripes and behind the flanking white stripes, the figure appeared as a single grey rectangle on a homogeneous black background that was partially occluded by white stripes. Similarly, when the grey elements were placed within the white stripes and behind the flanking black stripes, the figure appeared as a single grey rectangle on a uniform white background that was partially occluded by the black stripes. In terms of perceptual organization, this is tantamount to transforming the Munker-White display into a simultaneous brightness display. As noted above, this would imply that the magnitude of the Munker-White illusion should only be *as strong* as classical simultaneous contrast displays, since this perceptual organization is equivalent to displays that generate simultaneous contrast. For those observers experiencing this perceptual organization, this was what

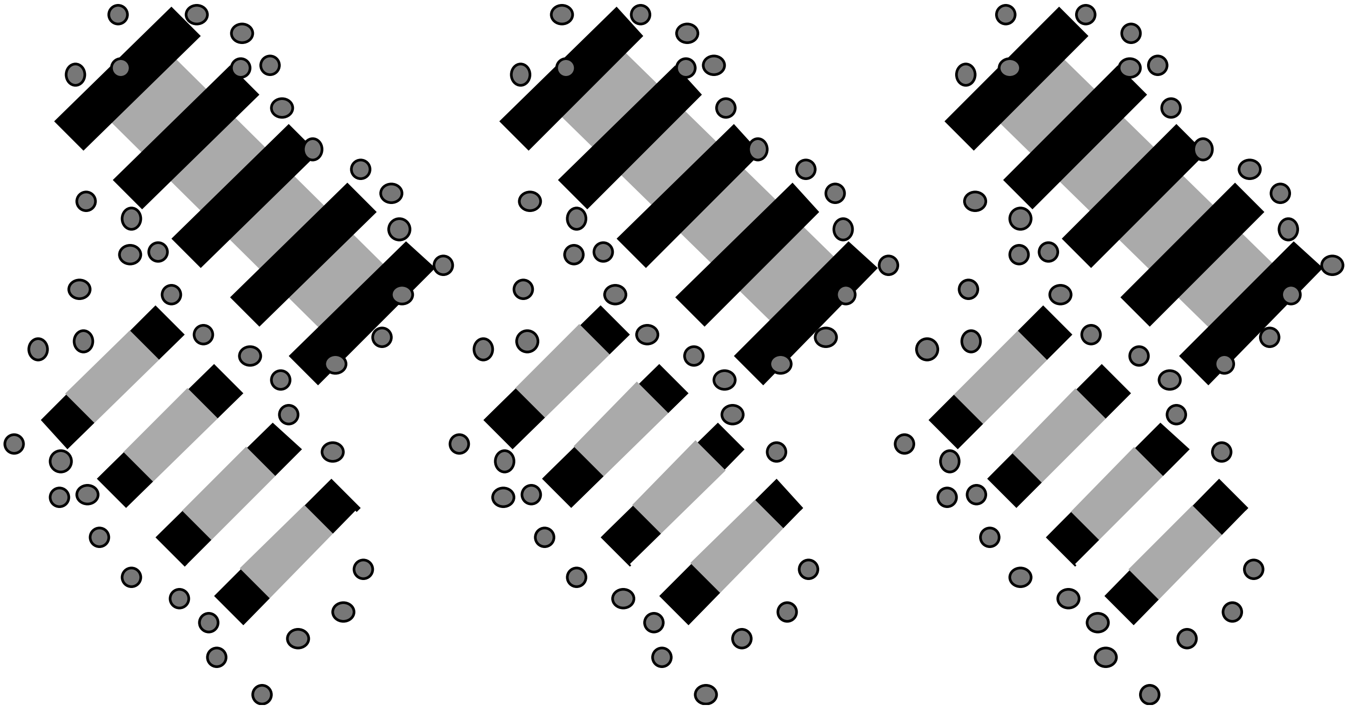


*Figure 11*

A stereo version of the Munker-White illusion, created by displacing the grey elements along the 45 degree bars in which they are embedded. The dots in this image serve as fusion locks that specify the epipolar lines. The illusion is much stronger and a clear percept of transparency is observed when the gray bars appear in front of the longer white and black bars. See text for details.

they reported.

However, the two observers that reported an enhancement of the Munker-White illusion when placed behind the rest of the display reported a very different perceptual organization. These observers described the black and white longer bars in which the grey bars were embedded as transparent (see Fig. 11), albeit less compellingly than the transparency experienced when the grey bars appeared in front. Nonetheless, this perceptual organization would be consistent with an enhancement of the difference between the two sets of grey bars. Consider the case in which the grey bars are embedded in the black stripes. In this case, the black stripes were perceived by these observers as two surfaces: a faint dark filter, overlying a grey bar on a dark background. In accordance with the scission hypothesis, if some of the darkness in the grey bar was attributed to the near (transparent) surface, then observers should experience the grey bar as being lighter than the monocular case. Similarly, when the grey bars embedded within the white stripes appeared behind the black flanking elements, the white stripes were reported to appear as a faint white film that overlaid grey bars on a white background. Again, since observers were asked to judge the color of the far grey bars, they should now appear darker than the monocular case, since some of the lightness of the grey target is now attributed to a near white filter, which is what these observers reported.[9]

Thus, I suggest that depth manipulations enhance the Munker-White illusion because this causes a stronger form of color scission than that observed monocularly. When the grey bars appear in front of the rest of the pattern, the grey bars in the black stripes appeared as a light filter through which a dark background is partially visible. Conversely, the grey in the white stripes appear as a dark filter that attenuates the light of a more distant, lighter surface. When the depth was inverted, achromatic color scission was still observed for some observers, albeit less strongly than the near case. This diminished strength of this scission is understandable, since the majority of the

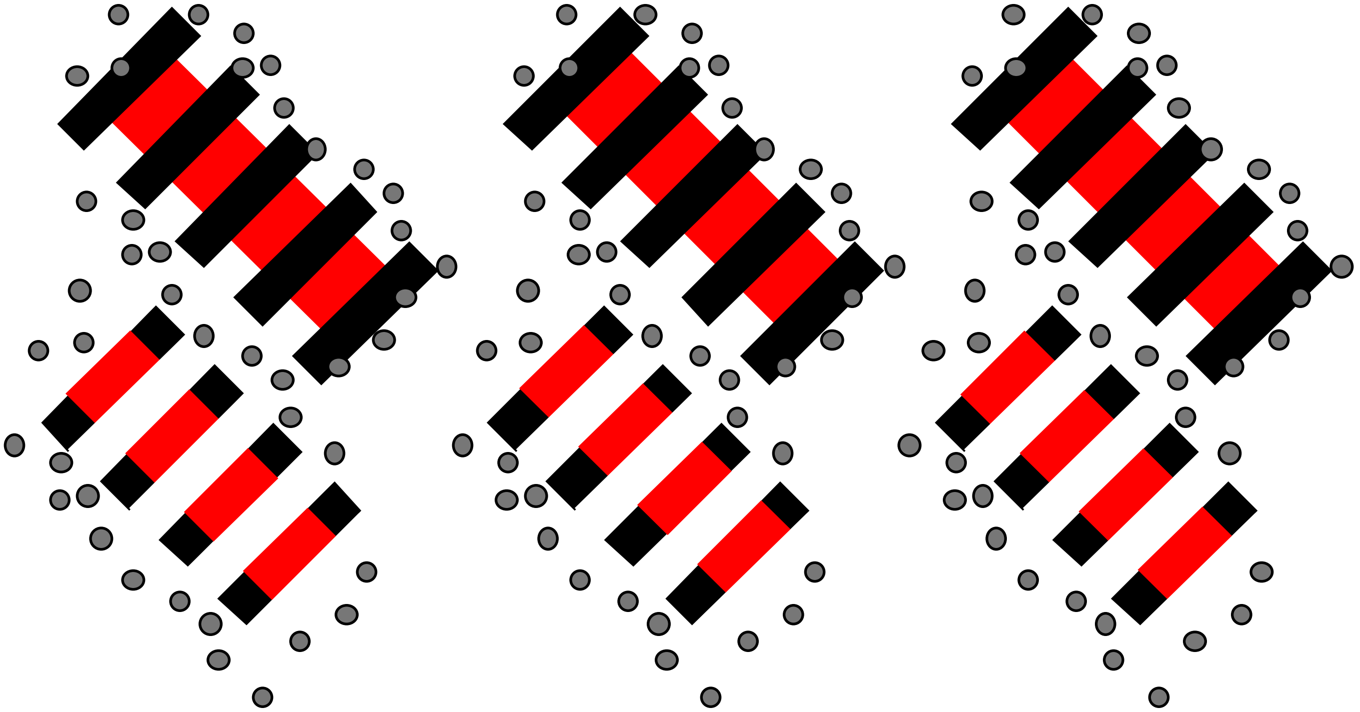


**Figure 12**

A chromatic variation of the Munker-White stereo display demonstrating the enhancement of the illusion with stereoscopic depth cues. See text for details.

luminance is still attributed to the background, and only some to the transparent surface. A similar result holds for the chromatic variants of the Munker-White illusion (see Fig. 12).

Although the preceding analysis has interpreted the enhancement of the Munker-White illusion as a consequence of scission, it should be noted that scission is not the only possible contributor to this illusion. I have argued that something other than contrast enhancement must be invoked to explain why the Munker-White illusion is stronger than simultaneous contrast, but it does not uniquely specify what that "something else" has to be. The argument here is that this process is scission, induced by collinear contours that preserve contrast polarity. However, there is nothing in the preceding arguments that necessitate that scission be the *only* process that contributes to the Munker-White illusion. Rather, it is possible that scission combines with a form of simultaneous contrast induced by the interpretation of the T-junctions as signalling occlusion. An explanation of this type would require a relatively modular organization of the mechanisms responsible for occlusion and scission computations that are subsequently combined in some kind of additive fashion. If such an organization existed, it would mean that theories that explained the Munker-White illusion with contrast enhancement mechanisms induced by occlusive T-junctions are not necessarily incorrect; they are merely *insufficient* unless supplemented by some other process (cf. Gilchrist et al., 1996).

A scission analysis can also be used to explain monocular brightness/lightness illusions such as Benary's triangles (see Fig. 13). In this case, the T-junction analysis would lead to the grey triangle within the black cross being decomposed into a near, light colored layer and a more distant dark surface, whereas the triangle placed in the white surround would be decomposed into a near dark surface that overlies a light background. Although Benary's illusion is weaker than the Munker-White illusion, this can be understood as a consequence of the fact that only one side of the triangle has T-junctions that would support the proposed mechanisms causing color scission. However, because Benary's illusion is not stronger than simultaneous contrast, we cannot rule out explanations of Benary's illusion that rely on an interpretation of T-junctions as a signature of occlusion, since this analysis would lead to the same direction of illusion as that observed with simultaneous contrast (see, e.g., Gilchrist et al., 1996).

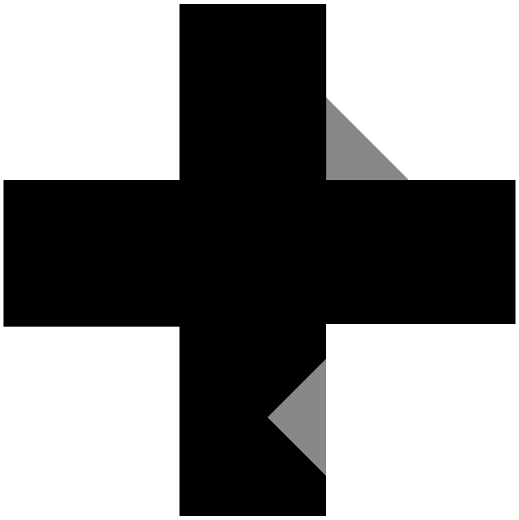


***Figure 13***

Benary's Cross. Note that the T-junction analysis also predicts the correct direction of this illusion.

In sum, it is argued that illusions such as the Munker-White's and Benary's are the consequence of mechanisms that sense a discontinuous change in the amount of contrast along aligned contours. By hypothesis, such mechanisms cause the lower contrast region to be split into two contributions: a near, transparent surface or illumination change, and a far surface visible behind this layer. To our knowledge, the analysis presented above can explain all of the brightness, lightness, and/or transparency illusions generated in displays containing polarity preserving monocular T-junctions. In the following section, this thesis is extending to displays generating I-junctions.

## Section 3: Color Scission in Monocular Images Containing I-junctions

The theory described above is easily extended to explain lightness illusions generated in images containing contours that are so thin that the orientation of their termination cannot be computed directly by even the smallest oriented receptive fields (cf. Grossberg & Mingolla, 1985). In this sense, these contour terminations do not "have" an orientation. Anderson & Julesz (1995) have previously termed such contour terminations "I-junctions," and I will retain this terminology here.

One of the most widely studied forms of lightness illusions is generated by images containing I-junctions, and has been termed neon-color spreading (Tuijl, 1975; Tuijl & Leeuwenberg, 1979; Tuijl & Weert, 1979; Ware, 1980; Redies & Spillman, 1981; Day, 1983; Redies et al., 1984; Grossberg & Mingolla, 1985; Nakayama et al., 1990; Watanabe & Takeichi, 1990). I contend that these illusions are identical to the Munker-White illusion, in the sense that the luminance and geometric conditions that give rise to them are identical. If this is true, and scission is the cause of the Munker-White illusion, then the same geometric and luminance conditions that were predictive of the Munker-White illusion should also predict when neon-color spreading should and should not occur.

There have been a number of spatial configurations used to evaluate the conditions that lead to neon-color spreading, but they share a number of properties. The most significant property for the present development is the presence of collinear segments that undergo an abrupt change in contrast. In this section, we will consider the Ehrenstein-shaped variants of these illusions, although the analysis that follows encompasses any variant of the illusion that contains collinear contour segments.

There are three general combinations of luminance that can be used to create these patterns, two of which are predicted to give rise to color scission within the context of the present theory. The two patterns that should give rise to color scission, and hence neon color spreading, are those that preserve contrast polarity. The pattern in which contrast polarity reverses should not give rise to neon color spreading, and hence, no lightness illusion or illusory transparency is predicted to occur.

An example of one of the Ehrenstein patterns that induces neon color spreading is shown in Fig. 14a. Note that this condition is identical to the region in the Munker-White illusion in which the grey bar is contained within the black stripe; the only differences is the size of the bars and the relative orientations of the multiple contour segments. In keeping with the preceding explanation of the Munker-White illusion, the scission thesis would attribute the "neon" quality of the color spreading as a consequence of treating the colored region embedded in the black stripe as a product of two distinct sources: an overlying light layer (the source of the neon color), and the continuation of the black bar behind this filter. This is also observed in the chromatic variant of the Munker-White illusion in Fig. 14a.

The other configuration predicted to give rise to color scission arises when the color of the large thin lines and the background are interchanged (see Fig. 14b). Note that this condition is identical to the region in the Munker-White illusion in which the grey bar is contained within the white stripe. Again, the proposed process responsible for this percept is scission, induced by the change in contrast along the contour's length, which is also observed in the chromatic variant of this illusion.

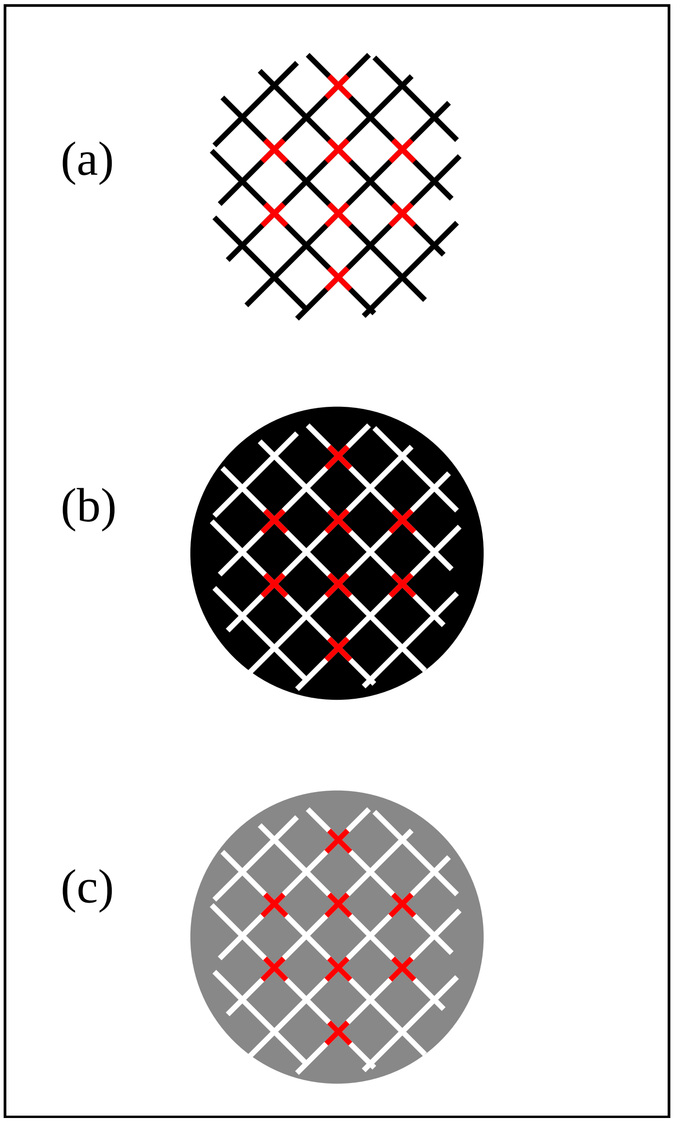


***Figure 14***

A series of neon color spreading displays. Note that color spreading does not occur when the background is the intermediate luminance (c).

Finally, the last condition that must be considered is when the aligned contours undergo a contrast reversal (see Fig. 14c). By hypothesis, we have argued that color scission will only occur if the contrast polarity of the aligned edges is the same. Hence, no color scission should occur in this display, and color spreading is not reported for these patterns. This is in accord with previous findings that have reported that neon color spreading only occurs when the target's luminance is between that of the background and the longer lines (van Tuijl & de Weert, 1979, a result that has a significant impact on at least one model of this phenomena [Grossberg & Mingolla, 1985; see the discussion of Bressan (1993b) and Grossberg & Mingolla's model in Section 5 below]). Within the present development, the primary reason *why* the target must be intermediate in luminance is that it is only these configurations that preserve the contrast polarity of aligned contour segments, which are presumed to be responsible for scission.

In sum, it has been argued that the conditions that support neon color spreading are identical to the conditions that generate large lightness differences in the Munker-White illusion. For all of these displays, we have argued the cause of the color spreading and the lightness illusions is chromatic and/or achromatic color scission induced by a discontinuous change in the magnitude of the luminance contrast of aligned contours. To this point, we have only considered the monocular conditions that give rise to the lightness and transparency illusions that are putatively a consequence of scission mechanisms. In the sections that follow, we extend this analysis to encompass the variations of these illusions generated stereoscopically.

## Section 4: Binocular junction geometry

In the previous sections of this paper, a general account of monocular illusions of lightness and transparency in images containing contour junctions was described, and a general rule of perceptual organization was proposed as an explanation of all of these illusions. In this section, I will describe how this rule can be extended to provide an understanding of binocular variants of these illusions. It may be of some historical value to note that the theory described above originated in the author's studies of binocular displays that generated percepts of illusory transparency. Within the context of the present paper, it is suggested that the primary role of stereopsis is to *reveal* and *enhance* the process of scission described above, rather than being the sole *cause* of scission. The problem, then, is to understand how and when stereoscopic viewing enhances the process of scission presumed to operate in monocular variants of these patterns.

To understand the power of the binocular contour junctions in generating percepts of illusory transparency, consider the stereopairs presented in Fig. 15. In the top stereopair, a clear percept of illusory transparency can be observed, and a faint tinge of red is seen to spread into the central black region (cf. Anderson & Julesz, 1995; Nakayama, Shimojo, & Ramachandran, 1990). Moreover, note that the disc shaped region behind the illusory transparent surface appears to have the same color as the white discs that the red transparent surface overlies, in accordance with our assumption that the more distant layer retains the same surface quality as the adjacent regions.

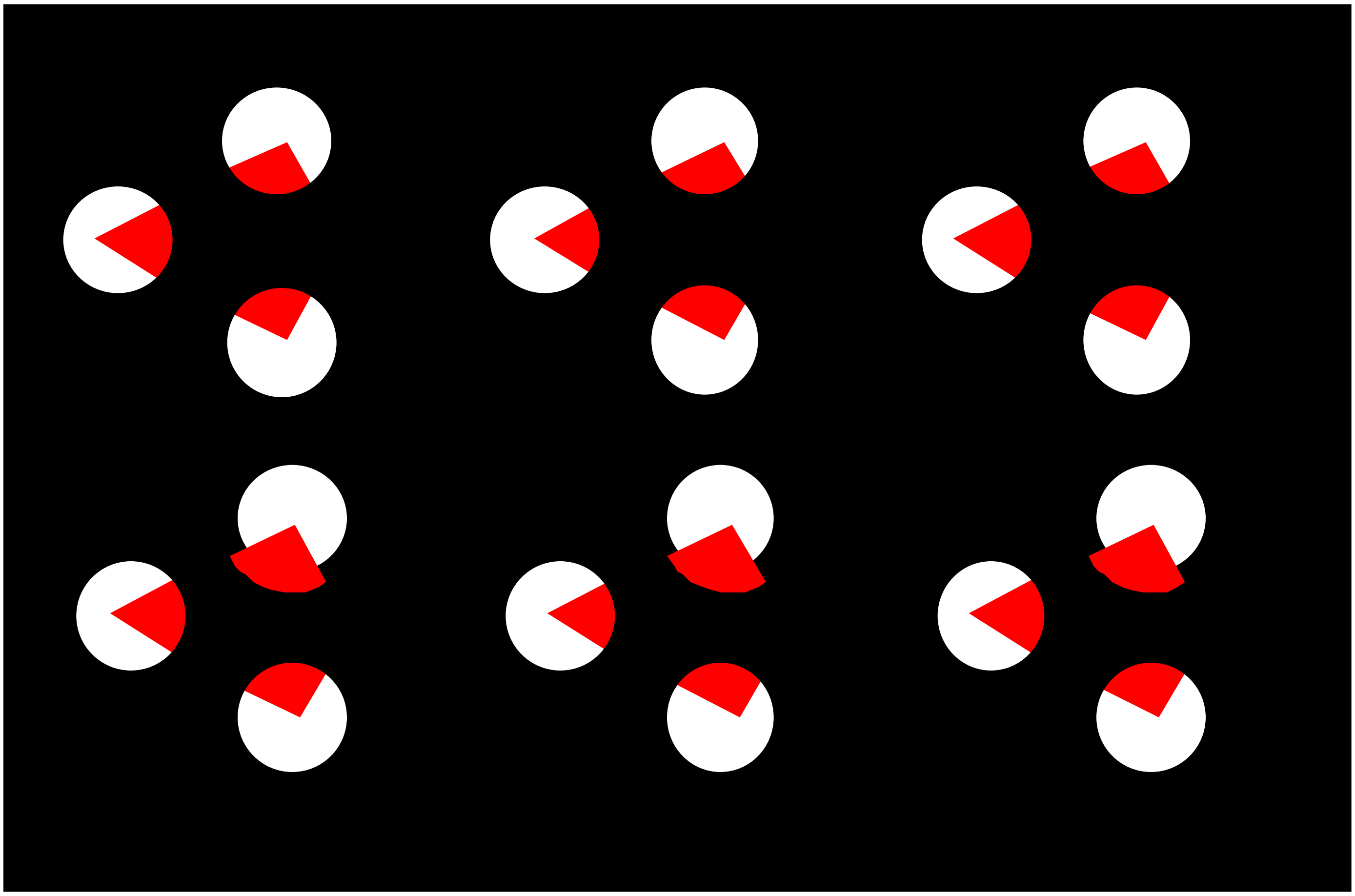


***Figure 15***

Stereopair demonstrating the importance of the local junction geometry in creating percepts of illusory transparency and neon color spreading. Cross fusers should fuse the left two images, divergers the right two. In the top figure, a clear percept of illusory transparency and neon color spreading is observed. In the bottom figure, the contours of the top right sector of the pattern are all at the same depth in front of the white elements and the black background, whereas the other two sectors are identical to the top. Note that the top right sector of the bottom stereopair appears as a simple, opaque surface patch that occludes the white disc and the black background, whereas the other two sectors retain an appearance of transparency. All of the red sectors are identical in color, yet appear extremely different.

However, in the bottom stereopair, a simple variant of the top stereopair is presented in which the junction geometry of the top right red sector has been slightly modified. In this region, all of the boundaries surrounding the red sector are given a common disparity so that they appear at the same depth. This leads to a change in the monocular junctions present in the two images. When the bottom stereopair is fused and this sector is compared with the other two, the importance of the local junction geometry can be immediately appreciated. The top right sector appears as an opaque red image patch that is occluding a white disc and a small portion of the black background. In contrast, the other two red sectors appear vividly transparent, and much of the lightness of these patches is perceived to belong to the presence of white discs behind the transparent surfaces. When viewing these images binocularly, it is difficult to believe that all of the red sectors are exactly the same color: the percept of transparency completely transforms the qualitative properties of the red sectors, effectively splitting a single color into two distinct layers.

The demonstrations presented in Fig. 15 motivate a closer look at the relationship of the junctions that give rise to such vivid percepts of illusory transparency. In these images, the red sectors terminate along the edges of the white circular discs, forming T-junctions. As noted above, T-junctions are usually indicators of occlusion, where the "stem" of the T is treated as occluded behind the "top" of the T. However, binocular parallax creates the possibility of placing the stem of the T-junction in front of the top of the T, which can give rise to three classes of illusory transparent surfaces (Anderson & Julesz, 1995). These three classes of percepts may be distinguished on the basis of their perceptual appearance, as well as by the kinds of stereoscopic T-junctions that give rise to these surfaces. I will describe each qualitative type of transparency in turn.

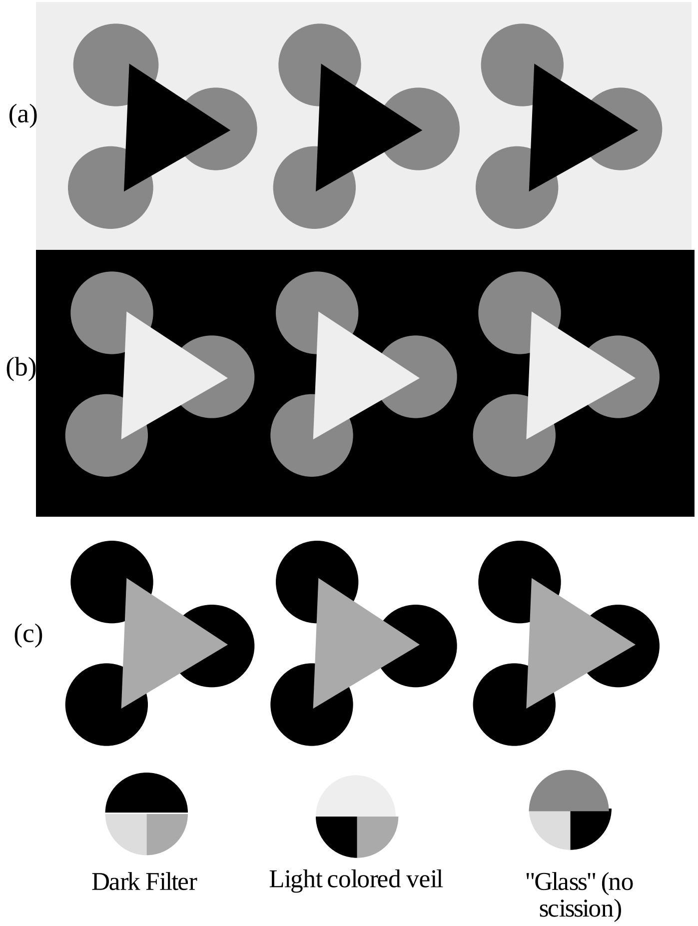


***Figure 16***

Stereopairs depicting the three types of illusory transparency that can be generated by binocular T-junctions. Cross fusers should fuse the left two images, divergers the right two. The top figure generates percepts of a dark filter that overlays a darker triangle and white background. The middle stereopairs generates percepts of a light mist or frost colored filter. The bottom stereopair generates percepts of illusory "glass," i.e., a perfectly clear transparent surface. The T-junctions at the bottom of the paper illustrate the luminance combinations that give rise to each of these percepts.

One class of T-junctions causes the formation of illusory surfaces that appear to attenuate a more distant bright surface, i.e., they appear as filters that attenuate the luminance of a brighter background. The stereo T-junctions that give rise to such percepts are shown in Fig. 16a. Note that in these figures, we are only considering the case in which the T-junction "stem" appears in front of the top of the T-junction; the reverse configuration typically gives rise to a simple percept of occlusion.

A second form of illusory transparency generates percepts of a light frost or mist colored veil that appears to overlay a dark background. In such images, the transparent surface appears as the primary contributor to the luminance in the region that it appears transparent. The stereo T-junctions that give rise to these transparency percepts are shown in Fig. 16b.

Finally, a third form of illusory transparency can be created that generates the appearance of glass, i.e., a transparent surface that appears perfectly clear. The stereo T-junctions that give rise to these percepts are shown in Fig. 16c.

What underlies these three forms of illusory stereoscopic transparency? A number of authors have argued that a variant of the Metelli conditions must be satisfied for the perception of transparency in images containing stereo T-junctions (Nakayama et al., 1990; Watanabe & Cavanagh, 1993). Roughly, this variant of Metelli's rule says that the intermediate luminance of the T-junction can be decomposed into two sources: a near, transparent surface, and a more distant surface visible through the transparent layer. This rule simply expresses the fact that an intermediate color may be created by a weighted combination of the other two luminances (light and dark), but neither the brightest or the darkest luminance in the scene can be created by a combination of the intermediate luminance with the other extreme.

Although the intermediate luminance rule correctly predicts a number of the stereo displays that generate illusory surfaces, there are a number of shortcomings of this rule when it is divorced from geometric considerations. Probably the most difficult problem with this thesis in understanding the percepts generated in the displays depicted in Fig. 16 is that *it does not explain why an intermediate luminance does not appear transparent when it is along the top of the T-junction*. Consider Fig 16c. Note that the grey triangle - the intermediate luminance - does not appear transparent when it appears in *front* of the discs and the background. Rather, it appears as an opaque occluding triangle, in front of amodally completed discs on a uniform background. Why doesn't this grey region split into two layers, if all that is required is that the transparent region be intermediate in luminance relative to the other two luminances? Clearly, something more than an intermediate luminance rule is needed to answer this question; some geometric properties must also be considered.

To date, all of the stereoscopic displays containing T-junctions that generate percepts of illusory transparent surfaces are accompanied by the formation of illusory contours. Therefore, any theory of illusory stereoscopic transparency should also consider the binocular conditions that lead to the formation of illusory contours. This is described in the following section. This section is included so that the present manuscript is relatively self-contained, and may be skipped by readers not interested in the formation of binocular illusory contours.

## 4.1: Illusory contour formation in binocular images

In general, there are two candidate mechanisms that have been proposed as the cause behind the formation of illusory contours (cf. Anderson & Julesz, 1995). In images containing contour intersections that form L-junctions (i.e., corners) or T-junctions, one proposed mechanism of illusory contour formation propagates the contour along its current trajectory (Grossberg & Mingolla, 1985; Kellman & Shipley, 1991; Field, Hayes, & Hess, 1993). Such approaches invoke *contour continuation* mechanisms to explain the interpolation of edges. However, in images containing only thin lines, illusory contours can form that are orthogonal (or nearly orthogonal) to the inducing edges. When line endings become sufficiently thin, their orientation will not be uniquely signaled by orientation selective mechanisms that compute orientation over a much larger image region. The term *end-cut* has been used to describe the process of generating illusory contours in the manner (Grossberg & Mingolla, 1985), a terminology that will be retained here.

In previous work, we demonstrated that contour terminations that contained unmatched segments in one of the two eyes formed end-cut illusory contours, generating percepts of illusory occluding surfaces (Anderson, 1994; Anderson & Julesz, 1995). This result revealed that contour continuation processes are not *necessary* for the formation of illusory contours in binocular vision, and hence, any theory that *requires* the presence of a contours along the direction that the illusory contour forms must be incorrect (cf. Kellman & Shipley, 1991). This is not to say, however, that contour continuation mechanisms do not play a role in illusory contour formation when such image features are present. However, there is currently no evidence that *uniquely* identifies these features as either necessary or sufficient for the formation of illusory contours. Here, I briefly review our previous theoretical analysis of binocular contour junctions, so that it may be extended and applied to understanding percepts of illusory transparency and brightness in stereoscopic images.

One of the simplest ways to motivate the utility of end-cut mechanisms can be appreciated by considering how binocular occlusion geometry generates contour junctions. One geometric fact of binocular contour junctions generated by one surface occluding another is that they generate both horizontal and vertical displacements in the two eyes (except for occluding contours that are perfectly horizontal relative to the line of sight; see Anderson, 1994; Anderson & Julesz, 1995; Malik, 1996). This can be appreciated immediately by performing the following exercise. Holding your head upright, place your left hand in front of your head and point your left index finger 45 degrees upwards (i.e., roughly towards 2 o'clock). Now place your right hand behind the left, and point your right index finger upwards and to the left -45 degrees. Arrange your hands so that the fingers appear to project an X-shaped image to each eye, while keeping the two fingers separated in depth with the right finer behind the left. Now alternately open and close your left and right eye, and observe how the V-shaped junctions appear to shift horizontally and vertically as the left and right eyes alternately open and close. Note also that the vertical shifts of these junctions alternatively occlude and reveal portions of the far finger. Following Belhumeur & Mumford (1992), we describe unmatched features on a partially occluded surface as *half-occluded.* If the visual system was capable of detecting these half-occluded features, then they could potentially provide information about the geometry of occlusion.

One simple method for recovering unmatchable features is to restrict matching to epipolar lines, i.e., one dimensional, horizontal "slices" of the two image planes. Epipolar constraints have been widely used in models of stereomatching (see, e.g., Dev, 1975; Jones & Malik, 1992; Marr & Poggio, 1976, 1979; Pollard, Mayhew, & Frisby, 1985; Sperling, 1970), but had not been previously used as a source of information about occlusion geometry until our work. Specifically,we have shown that the visual system appears to use an epipolar matching constraint to determine which features at contour junctions are half-occluded (Anderson, 1994; Anderson & Julesz, 1995), which in turn are used to recover the geometry of the occluding surfaces (through the formation of illusory contours). We previously found that a single half-occluded contour terminator can generate an end-cut illusory contour that appears both oriented and in stereoscopic depth (Anderson, 1994; Anderson & Julesz, 1995). Here, we will develop a generalization of this theory of illusory contour formation to include stereograms that generate percepts of transparency and neon color spreading.

We begin by summarizing the critical aspects of our previous theory of illusory contour formation in stereograms, so that we may then explain how these mechanisms may be used to understand how surface quality is attached to these contours. End-cut illusory contours can be generated by a single contour terminator, but only if a specific relationship holds between the eye-of-origin of the half-occluded feature and the relative position of the binocular portion of the contour (i.e., that portion of the contour visible to both eyes). The interested reader is directed to Anderson & Julesz (1995) for a complete derivation of these relationships. For the purposes of this paper, I will restrict myself to a single example to clarify the essential properties of binocular junction geometry. Consider a partially occluded contour oriented at 45 degrees, such as that depicted in the stereogram in Fig. 17a. When this figure is fused, a clear percept of a subjective contour arises, generating a near surface that appears to occlude a more distant surface. The horizontal line in Fig. 17b connects corresponding image junctions in the two half-images. Note that there is a portion of the contour oriented at 45 degrees that is only visible to the left eye. When these monocular terminators are attached to an appropriate fused contour segment (i.e., a segment that generates disparity signals), the monocular terminator will generate an illusory contour that appears both oriented and in depth (Anderson, 1994; Anderson & Julesz, 1995). However, not all monocular terminators are capable of generating illusory contours. For this example, an illusory contour will only form if a *right*-eye monocular terminator is attached to the bottom *left* of the contour, but not if a *left* eye monocular terminator is attached in the same position. In contrast, a illusory contour will form if a *left*-eye monocular terminator is attached to the upper right portion of the contour, but not if a *right*-eye monocular terminator is present in the same location. In both cases, these are the only configurations consistent with the unmatched terminator being part of a partially occluded surface, or equivalently, the only configurations in which the monocular terminator could be interpreted as the consequence of an occluding surface (this fact can be readily verified by the reader by using one hand as an occluder, and the other index finger as the occluded contour to determine the possible occlusion relationships). The important point of this example is that the formation of the illusory contour is dependent on *both* the eye-of-origin of the unmatched terminator *and* the geometric relationship of this terminator to the binocularly fused portion of the contour. We will see that this geometry is also critical in generating percepts of illusory transparency in stereopsis.

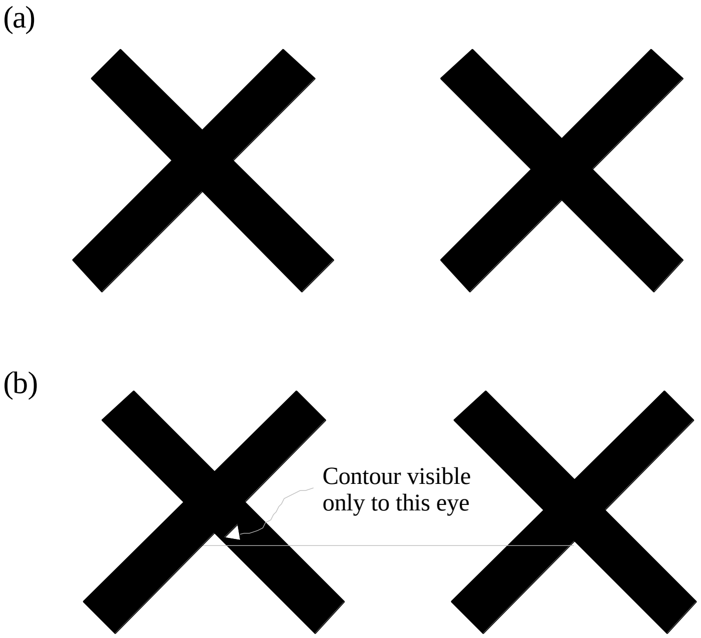


***Figure 17***

(A) A simple stereopair that generates a percept of an illusory occluding bar overlying a more distant bar. (B) Figure depicting the generation of unmatched contour segments under the assumption that the vertical shifts of the contour junctions are not matched (adapted from Anderson, 1994; and Anderson & Julesz, 1995).

Thus, when the visual system constructs illusory contours in stereograms consistent with the geometry of occlusion, the half-occluded features appear as a portion of the far surface and generate an illusory contour via an end-cut process. The primary difficulty that is encountered when attempting to generalize this mechanism to account for stereograms containing T-junctions that generate percepts of illusory transparency is that *there are no portions of the farther contour that are purely monocular in either eye in these images*. In other words, the far contour does not terminate in these displays, at least if "terminate" is taken to mean "disappear." This can be seen by studying Figures 16a,b. When fused, a percept of illusory transparent surfaces can be clearly observed, despite the fact that the far contours in these stereopairs -- the edges of the triangles -- have visible contrast along their entire length in both eyes. How is it possible to claim that end-cut responses are responsible for the formation of illusory contours when the far contour does not contain any features that are present in only one eye? Indeed, a simpler explanation would be to propose that the illusory contours are formed by a contour continuation mechanism that propagates the near contour (the T-stem) across the T-junction, as has been suggested previously (Kellman & Shipley, 1991). Thus, before an end-cut hypothesis is developed, we must first provide compelling

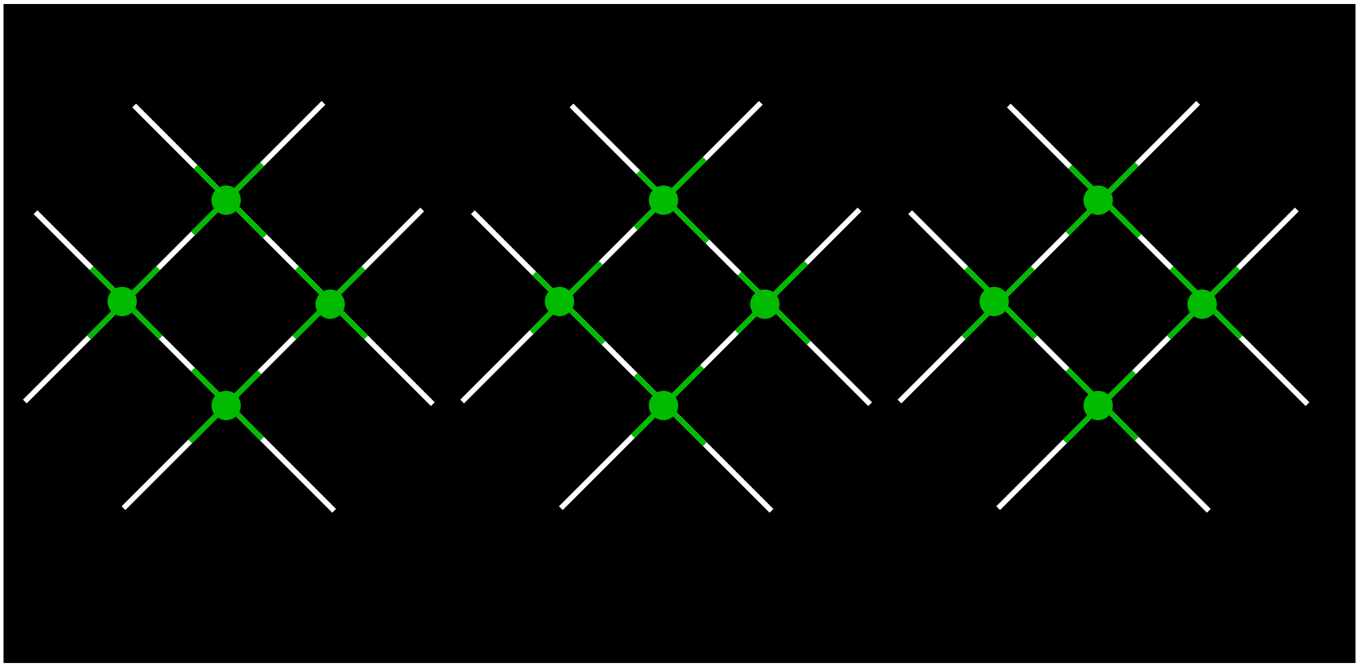


***Figure 18***

Stereo images that generate percepts of illusory transparent surfaces.  Cross fusers should fuse the left two images, divergers the right two.  These stereopairs were created by shifting the thin green contours along the 45 degree axis of the white lines.

evidence that there are end-cut type mechanisms that can give rise to illusory transparency in stereograms.  We must then explain how it is possible for an end-cut mechanism to work in displays such as those in Fig. 16 when there is no (apparent) contour termination that could drive it.

A natural place to search for the role of end-cut mechanisms in binocular vision would be to create stereoscopic variants of the stimuli that have been shown to create end-cut responses in monocular vision (Ehrenstein, 1941, 1947; Redies & Spillman, 1981; cf. Grossberg & Mingolla, 1985).  This led to the construction of the stereogram depicted in Fig. 18.  These stereograms were created by displacing the thin colored lines along the direction of the thin white lines that are oriented at ± 45 degrees.  Note that this means that these lines are displaced vertically and horizontally in the two eyes.  Under the assumption that only features along epipolar lines are matched, the vertical component of the colored contour's displacements will prevent the colored terminators from being matched, and hence, no disparity signals will be generated in these regions.  Moreover, these contours are sufficiently thin to rule out the possibility that there is any measurable orientation along the trajectory perpendicular to the thin line that could drive a contour continuation process. When these images are fused, observers uniformly report vivid percepts of transparent colored discs hovering in front of a uniform white grid.

Although the pattern presented in Fig. 18 generates clear illusory contours and percepts of transparency, there is a possible confound that must be ruled out before we can conclude that these contours are actually created by binocular end-cut mechanisms.  The problem is that the patterns in Fig. 18 generate monocular illusory contours, so it is possible that these contours input to conventional disparity mechanisms that, in turn, causes the percept of illusory transparency.  One method for eliminating the monocular illusory contours is to present luminance conditions that do not support neon color spreading.  As discussed in the previous sections, this occurs when the aligned contours undergo a contrast reversal.  As can be seen in Fig. 19, this stimulus does not generate illusory contours when viewed monocularly.  However, when this stereogram is fused, observers uniformly report the presence of illusory contours that appear to form perfectly clear transparent surfaces, such as glass.[10]

The preceding suggests two conclusions.  First, illusory transparency can be generated in images that do not contain horizontal disparities, and hence, disparity is not a

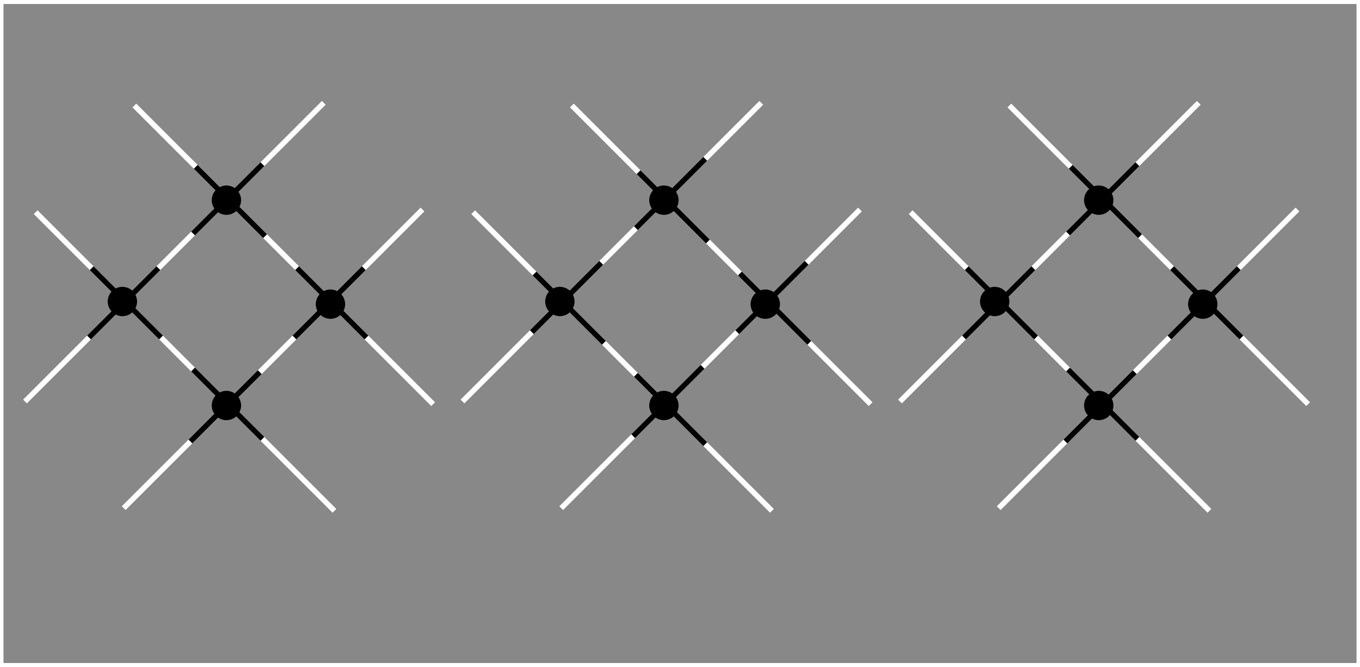


***Figure 19***

Same stereo images as those in Figure 18. The luminance conditions present in this figure do not support neon color spreading, yet observers report vivid stereoscopic illusory "glass" surfaces that appear in front of the figure. This demonstration reveals that the illusory contours present in Figure 18 can be synthesized by purely binocular mechanisms, and do not require the presence of monocularly visible color spreading signals to form.

necessary condition for specifying the relative depth of surfaces in stereograms generating percepts of illusory transparency. Second, illusory transparency can be generated in images composed only of thin lines that do not have any discernable orientation along the direction in which the illusory surface forms. This implies that contour continuation mechanisms are not necessary for the formation of illusory stereoscopic transparency (although, again, it does not rule out the possibility that contour continuation processes operate when they are present, or that such mechanisms are not used to group together the local end-cut inductions). Thus, some form of end-cut mechanisms must exist that are capable of creating percepts of illusory stereoscopic transparency.

With this conclusion in hand, the next problem is to understand how end-cut mechanisms can operate in displays containing T-junctions that do not appear to terminate (such as those in Fig. 16). The first step towards solving this problem is to explain what the analogue of half-occluded features would be in stereo images containing T-junctions. What is needed is a generalization of the notion of contour termination, and a generalization of "unpaired feature." In the simplest form, a contour termination may be regarded as occurring when the contrast of a contour goes from some finite value to zero, that is, when the contour literally ends in the image. However, the concept of "contour termination" does not have to be restricted to such extreme conditions; we can also consider a discontinuous change in *magnitude* of contrast as a form of contour termination. Indeed, if Figs. 16a and 16b are examined closely, it is evident that there is a discontinuity in the *amount* of contrast that arises along the contours of the triangles in these patterns. Note that although the triangle has contrast along its entire length, there is nonetheless an interocular difference in the *magnitude* of contrast at the junctions where the illusory contours form. This means that one eye has a higher contour contrast in this image region than the other eye, which is a step in the right direction of generalizing the notion of contour termination.

Although promising, generalizing the concept of contour termination to include discontinuities in the magnitude of contrast is still not sufficient to explain how illusory contours form. A substantial problem still

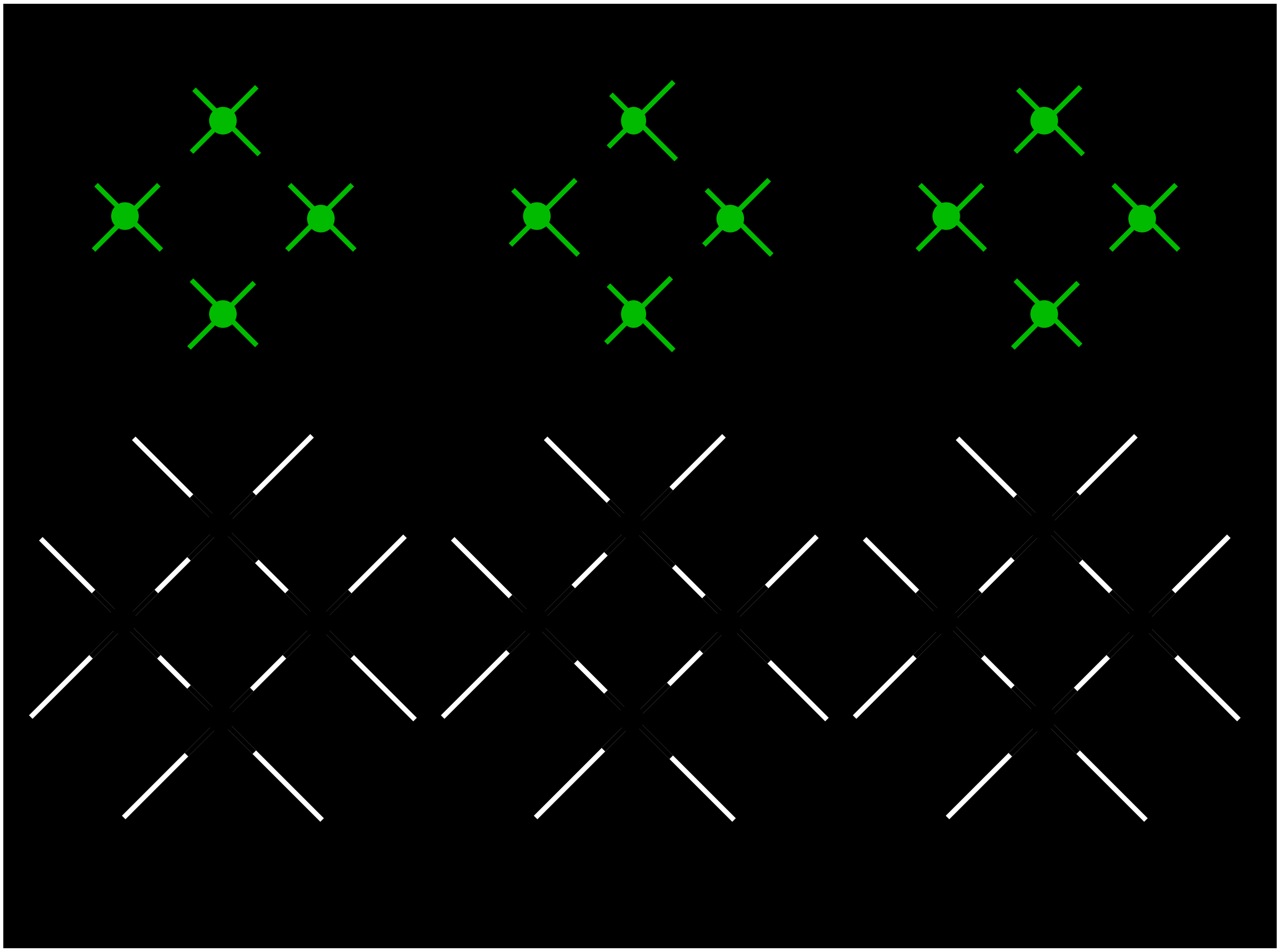


***Figure 20***

Figure in which the components of Fig. 19 are broken into separate elements. In the top row, this causes the formation of illusory holes in front of and surrounding the green lines (cross fusing the left two images, or diverging the right two). In the bottom figure, these patterns give rise to occluding discs, whose contours are in exactly the same position as the holes in the top of the figure. See text for details.

remains: we must now present a principled account of how the visual system determines which side of the binocular portion of the contour should be interpreted as *owning* this contrast mismatch. To understand the meaning of this statement, let us go back to the example of the 45 degree contours that generated percepts of illusory transparency and occlusion. When we were considering a 45 degree contour in the images generating illusory occluding contours, we found that a right-eye monocular terminator would only generate an illusory contour if it was attached to the bottom left of the contour (again, this can be understood by using one hand as an occluder, and the index finger of the other hand as the contour to determine the valid occlusion relationships). If the right-eye monocular feature was attached to the upper right of the fused contour segment, no illusory contour would form. Hence, the extension of our analysis to displays generating transparency critically depends on correctly assigning the interocular contrast differences to the appropriate binocular contour segment. How can this be done?

Previously, Anderson & Julesz (1995) suggested that the visual system treats interocular contrast mismatches in a manner similar to the contrast mismatch created by the purely monocular features generated by occlusion. Our linking hypothesis was that the unmatched terminators generated by occluding contours could be regarded as a form of "extra" contrast in one of the two eyes, analogous to the additional contrast generated by unmatched features generated by occluding surfaces. We therefore suggested that the problem was to determine which portion of the binocularly fused contour should be viewed as the binocular segment that "owned" this extra contrast. Our only guiding principle was that this assignment should yield the same pattern of matchable and unmatchable features that we discovered generated illusory occluding contours. This led to the conclusion that the extra contrast should be assigned to the binocular segment that was highest in contrast. This argument is somewhat dissatisfying, however, since it involves a circularity: The extra interocular contrast should be assigned in such away that it maintains the veracity of the theory we were proposing. As we will see, this led to the correct assignment, but there is a more principled way to motivate it.

My revised explanation builds on the following observation: in stereo displays such as those depicted in Figs. 18 and 19, there are two perspectives that may be taken on the contrast terminators present in these images. Consider a single element in the displays presented in Fig. 18. To remain consistent with our previous example, consider a green contour oriented along the 45 degree axis that is displaced up and to the right in the left eye. The top right terminator of this contour may be viewed as *either* an unpaired *green* feature present in the *left* eye, *or* an unpaired *white* terminator in the *right* eye. In other words, the determination of which eye contains the "extra" feature is a relative decision that depends on the contour under consideration (white or green). The critical insight is that *both interpretations lead to the synthesis of an illusory contour appearing in front of and oriented (roughly) perpendicular to their respective terminator*. In other words, both assignments generate the same illusory contour; the only difference is in the border ownership of the two patterns. Recall that the conditions that must be satisfied for such contours to form in occluding displays is a left-eye monocular terminator attached to the *top* of a 45 degree fused contour, or a right-eye monocular terminator attached to the *bottom* of a 45 degree fused contour. This is exactly the conditions that would be met by the method of assignment just described.

The preceding arguments can be easily understood by considering what happens to the display in Fig. 18 in the limit in which the white contours become the same color as the background (i.e., only the green contours remain visible). This stimulus would generate the appearance of a hole in front of and surrounding the small green contours (see Fig. 20, top). Similarly, consider what happens when the white contours are retained, and the green contours become the color of the background. This stimulus generates the appearance of small discs that occlude the ends of the white contours. The relevant point is that the contours in the display generating the illusory discs (Fig. 20, bottom) occur in *exactly the same position* as the display generating the illusory holes (Fig. 20 top). Thus, there is no need to "decide" which portion of the contour a given monocular terminator should be attached to; the monocular features are attached to those portions of the contour that have the same contrast in the two eyes (or more generally, that are most similar in contrast). Both assignments lead to the synthesis of the same illusory contours.

In sum, it was argued that end-cut mechanisms that respond to an interocular difference in the amount of contrast along the length of a contour play a strong role in the synthesis of illusory contours in stereoscopic displays generating percepts of illusory transparency. Anderson & Julesz (1995) previously made a similar suggestion, but did not have an unambiguous demonstration to support it, or a sufficiently principled argument that would allow for the proposed link of these end-cut mechanisms to percepts of transparency. Our previous work only demonstrated that interocular contrast differences support the formation of end-cut illusory contours, but these displays did not give rise to any clear sense of transparency (see also Kumar, 1995, for similar demonstrations). Therefore, we could only suggest the possibility that a common mechanism subserved both illusory transparency and illusory occlusion in stereopsis, but we could not convincingly demonstrate it. The new contribution of the demonstrations presented in Figs. 18 & 19 is that they reveal that end-cut responses can generate percepts of illusory stereoscopic transparency.

### 4.2: Attachment of surface quality to binocular contours

With our analysis of illusory contour formation in hand, the question that we must now address is how surface quality is partitioned between the real and illusory contours in these stereo displays. There are two cases to consider: T-junctions and I-junctions, both of which generate illusory contours. We have already demonstrated that the binocular junctions that generate percepts of illusory transparency are those in which the aligned contours preserve contrast polarity. As before, "aligned contours" refers to the two contour segments along the top of the T-junction, or the two thin contour segments that form an I-junction (see Figs. 18 and 19). We have also shown that a change in the magnitude of contrast can serve as a form of unpaired feature that can generate illusory contours. The only problem that remains is to determine how surface quality is attributed to this edge. To this end, all that is needed is the recognition that the side that appears transparent is the side of the aligned contour segment that is lower in (luminance) contrast. The preceding sections of this paper have suggested that the lower contrast region of aligned contours are decomposed into two causal layers. The role of stereopsis is to reveal and enhance this scission by introducing explicit depth signals at the contour discontinuities.

It is interesting to note that the visual system seems to encode transparent surfaces into two broad categories: those that attenuate the reflectance of a far surface (such as that which occurs when viewing a portion of a scene through a neutral density filter); and those that add luminance to a more distant, dark surface (such as might occur when looking through a window that is partially covered by a uniform layer of frost). The former putatively correspond to the apparent decrements in lightness observed when the small bars in the Munker-White display are embedded in the white stripes (since some of the luminance of the small bars is putatively treated as belonging to the extending white bars that continue behind these regions), and the latter corresponds to the neon or frost-like quality observed in the Ehrenstein variants of neon-color spreading (since some of the luminance of the small contour segments is putatively treated as belonging to the more distant dark contour).

## Section 5: Possible Mechanisms and Relationship to Previous Work

In the preceding sections of this paper, I have argued that the formation of illusory transparent surfaces involves two processes, both of which are driven by discontinuities along aligned contour segments. First, uniocular or interocular differences in contrast that arise from discontinuities of aligned contours are used to synthesize illusory contours that form (roughly) perpendicular to the inducing contour. Second, luminance relationships along aligned contours that undergo a discontinuous change in magnitude -- but not polarity -- induce a scission of the lower contrast region into multiple layers. When both processes are present, the result is the formation of an illusory transparent surface, or an illumination and lightness change.

The theory developed in the present paper relies on the geometric structure and luminance relationships that arise at contour junctions. This would seem to suggest the need for a rather large family of "junction detectors" that could classify the kind of junction present, and then interpret the possible causes of the junction. However, it was argued that the primary determinant of the brightness and transparency percepts created by I-, T-, and X- junctions are the monocular polarity relationships of the aligned contours in the junctions, and for stereoscopic images, the binocular shifts in the positions of these contours in the two eyes. This suggests that the mechanisms that compute these properties must be sensitive to contour orientation, contrast polarity, and the binocular positions of contours and contour terminators. Mechanisms with many of these properties are known to exist at the earliest stages of cortical processing, i.e., in area V1. Thus, no new "detectors" are required in the present account; all that is needed is a new kind of functional interaction between oriented units tuned to contour orientation, interocular position, and contrast polarity relationships.

A number of authors have described ideas related to those presented herein to account for a variety of lightness and transparency illusions (Bressan, 1993a,b; Bressan, 1995; Watanabe & Cavanagh, 1993; Redies et al., 1984). Redies et al. (1984) suggested that illusions such as neon-color spreading may be the consequence of end-stopped cells in V1 that (putatively) signal the ends of contours, similar to our assertion that end-cut responses play a critical role in these illusions. Probably the closest analysis of neon-color spreading to that developed here was by Bressan (1993a,b). Bressan argued that neon color spreading was observed whenever the figural and luminance conditions were suggestive of a transparent surface. However, she also reported the existence of what she termed neon-color spreading in displays that do not satisfy the conditions for transparency (i.e., conditions that would never be satisfied for physical transparency; see Bressan, 1993b; see also Beck et al., 1984). For the present analysis, the most problematic case she discussed was when the background was intermediate in luminance between the target and background lines. Within the theory described here, this condition should not give rise to scission, since the aligned contours in these displays undergo contrast reversals. To date, neon color spreading has not been reported for Ehrenstein or lattice patterns when the aligned contours undergo contrast reversals (see, e.g., van Tuijl & de Weert, 1979). Why, then, did Bressan (1993) and Beck et al. (1984) report instances of phenomenal transparency in stimuli that are physically inconsistent with the occurrence of transparency?

Although there is no conclusive answer that can be given to this question given the existence of conflicting data, Bressan did suggest a possible reason why it would be possible to observe something like transparency and/or neon color spreading in the displays used in her studies. The patterns reported by Bressan that violate the luminance conditions for transparency were Varin patterns, and she notes that these figures "..suggest interposition so strongly that a subjective figure is seen anyway..(1993b, p. 61)." A similar observation holds for the exceptions reported by Beck et al. (1984), which served as one of the motivations for the experiments reported by Bressan. However, Bressan notes that such examples of neon color spreading "..have, phenomenally, little in common with the familiar neon color spreading phenomena. One tends to look for vagueness and luminescence and a misty veil, and one may find neither in figures in which the luminance of the target lines is not in the middle....(1993b, p. 62)." Thus, to the extent that Bressan's results provide examples of neon color spreading that are not explicable within the present framework, these results have only been reported for a subset of patterns that have been studied as variants of neon color spreading (Varin figures), and there are clear phenomenological differences between these patterns and those previously termed "neon color spreading." It will remain for future research to determine why observers can report the appearance of transparent surfaces in image conditions that could not have been generated by a real transparent surface.

Grossberg (1994) and Grossberg and Mingolla (1985) developed a theory of neon color spreading that shares a number of properties with the account developed here. In general, Grossberg and Grossberg & Mingolla model is composed of two essential subcomponents. First, a boundary contour system (or BCS) senses and pools oriented contrast of opposite contrast polarity, and through local interactions with other local oriented contrast sensitive cells synthesizes a boundary web. The boundary web serves as a "skeleton" that is filled-in by a feature contour system, or FCS, which retains information about direction of contrast. Together, the BCS and FCS comprise the major component of their model that is used to explain effects such as neon color spreading.

To better understand the relationship between Grossberg and Mingolla's model and that described here, let us consider their account of neon color-spreading. First, the contrast difference between the thin contours of the background lines and the target lines causes the lower contrast, aligned segments of the target to be inhibited by the higher contrast line segments in which the targets are embedded. This inhibition then causes the orthogonally oriented cells at the boundary of the target and background lines to gain a competitive advantage, causeing the formation of end-cuts. The end-cut outputs are then fed into a cooperative feedback loop that dynamically adjusts that outputs of this competitive stage and synthesizes a global illusory contour that groups the local end-cut response together into a roughly circular boundary. Once this closed contour is formed, the FCS initiates a color spreading signal which propagates the color within the target lines up to the boundaries of the illusory circular contour generated by the BCS (see Fig. 21).

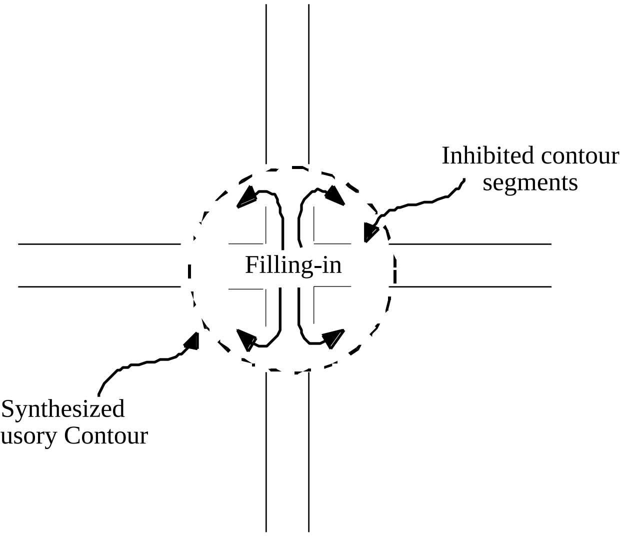


***Figure 21***

A depiction of Grossberg & Mingolla's (1985) model of neon color spreading. In this model, aligned contour segments with similar orientations are assumed to compete. If there is a difference in the contrast magnitude of the contour segments, then the larger contrast region is assumed to have a competitive advantage, and the adjacent, lower contrast contour segments are inhibited. This inhibition putatively imbued orthogonal orientations with a competitive advantage (due to disinhibition), creating roughly orthogonal end cut responses that group to form a coherent boundary (dashed lines). The inhibition of the aligned, lower contrast contours is assumed to allow color to flow out of the target limbs in the Ehrenstein figure up to the edges of the illusory circular boundary. A critical feature of this model is that the competition which suppresses aligned contours of the same orientation depends only on the *sign* of contrast, not on the magnitude of contrast. Thus, neon color spreading should occur whenever there is a significant contrast difference between the target contour segments, and the background contours.

Thus, the model of neon color spreading advocated by Grossberg & Mingolla assumes that neon color spreading is caused by the difference in the magnitude in contrast between the background and target contours. Although this model goes further towards articulating a clear mechanism for neon color spreading than that described herein, there seems to be a problem with the Grossberg & Mingolla model in its treatment of contrast polarity, or more specifically, in its failure to treat contrast polarity as a critical image property. In both the Ehrenstein and lattice variants of neon color displays, color spreading only occurs when the luminance of the target lines lies between the luminance of the background and the background lines (van Tuijl & de Weert, 1979). One of the fundamental properties of Grossberg & Mingolla's model that has survived its evolving incarnations is the assumption that "..the boundary process is sensitive to the *amount* of contrast, even though it is insensitive to the *direction* of contrast (1985, p. 182, emphasis theirs)." Thus, the formation of boundary contours is indifferent to the direction of contrast along the aligned contours; the only property that is relevant to the inhibition of the target segments and the release of the color within the targets is the *amount* of contrast difference. The problem with this explanation of neon color spreading is that it is possible to create a neon stimulus wherein the target and external lines undergo a polarity reversal, but in which the target's contrast magnitude is smaller than that of the external lines (see Fig. 22). Within the framework of Grossberg & Mingolla, this should induce neon color spreading, since it is only the magnitude of the contrast difference between the external lines and the targets that leads to the synthesis of the illusory boundary contours. Thus, Grossberg & Mingolla's model predicts that neon color spreading should occur in these conditions, but it does not. This is also true of all known stereoscopic variants of these patterns (Anderson & Julesz, 1995; Nakayama et al., 1990).[11]

A second difference between Grossberg's (1994) and Grossberg & Mingolla's (1985) theory and that described here is that there is currently no mechanism available to generate illusory stereoscopic contours generated by unpaired contour terminators (cf. Anderson, 1994; and Anderson & Julesz, 1995). We have demonstrated here and previously that vertical interocular displacements of contours can generate end-cut responses that synthesize illusory contours, putatively because these vertical displacements are treated as monocular (i.e., unmatched). The theory of Grossberg (1994) has attempted to localize unpaired features at the same depth as the binocularly fused portion of the contour; it does not use these features to synthesize illusory contours. Hence, the approach outlined in the present paper currently encompasses a number of displays that are not explainable within the approach described by these Grossberg & Mingolla.

In fairness, it should also be noted that an extension of Grossberg & Mingolla's model described by Bressan (1995) has one significant advantage over the perspective described here. Specifically, Bressan (1995) has argued that a variant of the Grossberg & Mingolla model may provide a sufficient account of chromatic neon color spreading reported by Ejima, Redies, Takahashi, & Akita, (1984). In this paper, I have limited the term "contrast" to refer to luminance differences under the assumption (and informal observations) that achromatic contrast is the primary determinant of the illusions described above. Bressan's extension of the Grossberg and Mingolla model to

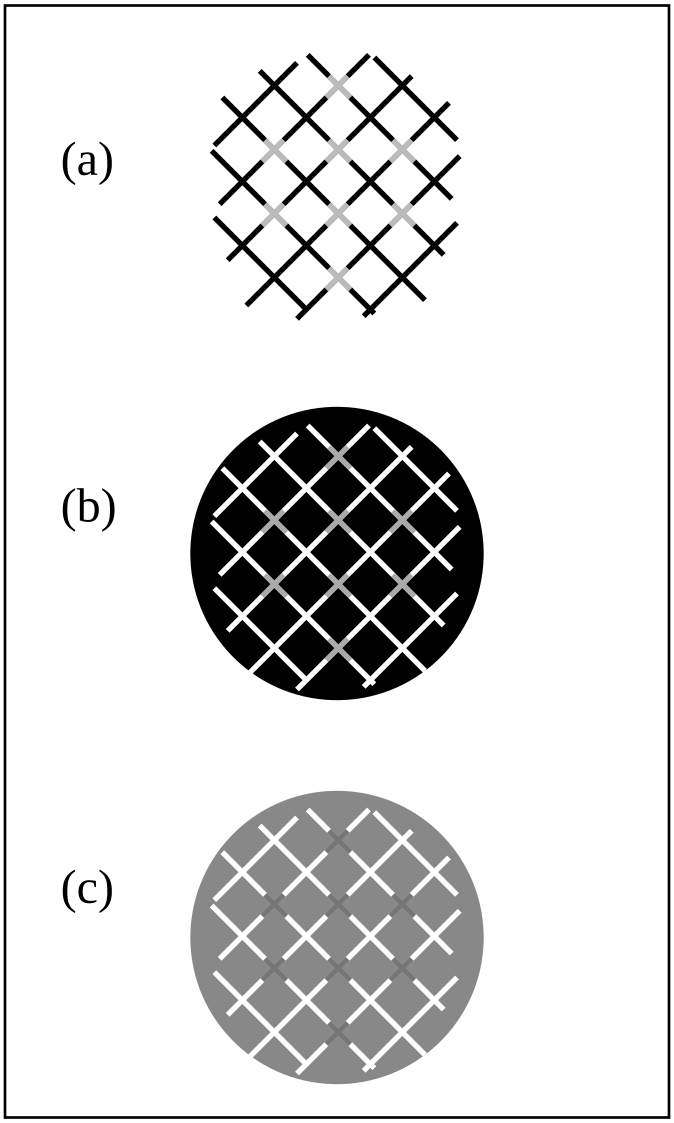


***Figure 22***

Figure demonstrating that neon color spreading does not occur for low contrast targets when the aligned contours undergo polarity reversals, in contrast to the prediction of the Grossberg & Mingolla model. See text for details.

handle complementary color induction which occurs in some chromatic variants of neon color spreading has not been addressed in the present theory, and remains beyond the scope of this paper.

Finally, Watanabe & Cavanagh (1993) have also suggested that a variety of displays containing T-junctions can behave as if they were X-junctions generating percepts of transparency, and termed such T-junctions "implicit X-junctions." They suggested that neon color spreading could also be understood from this perspective, and noticed that a scaled-up variant of the Ehrenstein figure would generate T-junctions. Although this account shares a number of properties with that developed here, Watanabe & Cavanagh's account of neon color spreading seems somewhat improbable, since neon color spreading is most vivid for very thin contours that do not have contrast along the orientation in which the illusory contour forms, and T-junctions would be clearest for thicker inducing elements. In other words, neon color spreading is strongest when the contour junctions in the images appear the *least* like T-junctions, not when they contain clearly visible T-junctions. In this paper, I have suggested that the percept of transparency generated by X-, T-, and I- junctions all depend on a single image property: namely, the contrast relationships of the aligned contours that arise at the image junctions. From this perspective, the stem of the T-junction (the contour that crosses the contrast change of the aligned contour segments) is not necessary for color scission to occur. The only requirement is a contrast change along the aligned contours in the images. In keeping with Watanabe & Cavanagh, the present account assumes an equivalence between I-junctions in neon-color spreading displays and illusory transparency induced by T-junctions. However, this equivalence is based on the presumption that the critical property in T- and X-junctions is their similarity to I-junctions, not conversely. Thus, although there are similarities in the set of stimuli that Watanabe & Cavanagh and I treat as equivalent, we assume that different image properties are determining this equivalence, and hence, that different kinds of mechanisms underlie these phenomena.

## Section 6: Limitations of the Theory and Future Directions

In this paper, a qualitative theory of monocular and stereoscopic illusions of transparency and lightness was presented that relies on the concepts of scission and end-cut illusory contours. The majority of previous theories of these illusions have attempted to interpret these effects as the consequence of mechanisms that enhance the contrast of image structure (see Pessoa et al., 1996; and Pessoa, 1996, for a recent review), although there are a few notable exceptions (Adelson, 1993; Gilchrist, 1977 ). The computational rationale for such mechanisms is to amplify luminance differences and thereby enhance the visibility of image structure. In the approach described here, I take a very different tact on the computational goals that are assumed to underlie many lightness and transparency illusions. Specifically, I have argued that the primary goal of visual processing is not to simply amplify *image* structure, but rather, to *decompose* image structure into the underlying causes of the images. The mechanisms underlying of scission -- producing a layered representation -- is just one process putatively used to decompose images into the variety of causes that were generated by the projection of surface properties onto our retinas.

In this paper, I have articulated a simple rule that putatively underlies the decomposition or scission of a given image region into multiple layers. The focus of this theory has been on the local structure that must be present for a decomposition into multiple layers, namely, aligned contours that contain abrupt discontinuities that preserve contrast polarity. In all of the patterns considered here, there was always more than one junction present along the aligned contours (typically two, one on either side of the target). Thus, the local junction rule outlined above should be understood as entailing a requirement that no polarity reversals occur over the entire length of the aligned contours along a target region. In this paper, my goal has been to demonstrate that scission is capable of providing a principled explanation of a large number of different lightness and transparency illusions that share qualitative photometric and geometric properties. A more detailed analysis of the geometric determinants of scission will be addressed in a subsequent paper.

At present, the theory described here only makes predictions about the sign and the ordinal magnitude of illusion strength in images containing X-, T-, and I-junctions that contain aligned contours which preserve contrast polarity. However, it is well known that striking brightness illusions can be created in patterns containing T-junctions that do not preserve contrast polarity, such as that generated by grating induction (McCourt, 1982). Unlike the illusions described in the present paper, the illusions described by McCourt and his colleagues generate inhomogeneous brightness inductions, similar to those of Mach bands. In its present form, the theory described here makes no prediction about the consequence of images containing aligned contours that undergo contrast reversals (other than to assert that a homogeneous scission should not occur), and therefore remains mute in addressing phenomena such as grating induction. In the future, it is hoped that the present framework can be extended to handle spatially inhomogeneous illusions, lightness and transparency illusions created in images that do not contain contour junctions, as well as brightness illusions induced by second-order stimuli (Chubb, Sperling, & Solomon, 1989; Solomon, Sperling, & Chubb, 1993; Lu & Sperling, 1996).

**Acknowledgements**

I would like to thank A. Gilchrist, E. Mingolla, and H. Sedgwick for their penetrating reviews of an earlier version of this manuscript, and E. Adelson, D. Somers, and Y.Weiss for their thoughtful comments on many of the ideas described in this paper. This word was supported by NIH grant EY0D11487-01 to B. Anderson.

## References

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

Adelson, E.H., Anandan, P. (1990) Ordinal characteristics of transparency. AAAI-90 Workshop on Qualitative Vision, July 29, 1990, Boston, MA.

Anderson, B.L. (1994) The role of partial occlusion in stereopsis. *Nature*, **367**, 365-368.

Anderson, B.L. (1995) New phenomena in surface interpolation. *Investigative Ophthalmology & Visual Science Suppl.*, 1732.

Anderson, B.L. (1996) Partitioning contrast into distinct 3D surfaces. *Investigative Ophthalmology & Visual Science Suppl.*, 1328.

Anderson, B.L., Julesz, B. (1995) A theoretical analysis of illusory contour formation in stereopsis. *Psychological Review*, **102**, 705-743.

Barrow, H.G., Tenenbaum, J.M. (1978) Recovering intrinsic scene characteristics from images. In Hanson, A.R., and Riseman, E.M. (Eds.), *Computer Vision Systems*, 3-26.

Beck, J. (1985) Perception of transparency in man and machine. In Human and Machine Vision II, A. Rosenfeld (ed.), New York, Academic Press.

Beck, J., Prazdny, K., and Ivry, R. (1984) The perception of transparency with achromatic colors. *Perception & Psychophysics*, **35**, 407-422.

Belhumeur, P.N., & Mumford, D. (1992) A bayesian treatment of the stereo correspondence problem using half-occluded regions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Urbana, IL.

Bressan, P. (1993). Neon color spreading with and without its figural prerequisites. *Perception*, **22**, 353-361.

Bressan, P. (1993). Revisitation of the luminance conditions for the occurrence of the achromatic neon color spreading illusion. *Perception & Psychophysics*, **54**, 55-64.

Bressan, P. (1995). A closer look at the dependence of neon color spreading on wavelength and illuminance. *Vision Research*, **35**, 375-379.

Cavanagh, P., & Leclerc, Y. (1989). Shape from shadows. *Journal of Experimental Psychology: Human Perception and Performance*, **15**, 3-27.

Chubb, C., Sperling, G. and Solomon, J. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences USA*, **86**, 9631-9635.

Clowes, M. B. (1971) On seeing things. *Artificial Intelligence*, **2**, 79-116.

Day, R.H. (1983) Neon color spreading, partially delineated borders, and the formation of illusory contours. *Perception & Psychophysics*, **34**, 488-490.

Dev, P. (1975) Perception of depth surfaces in random-dot stereograms: A neural model. *Int. J. Man-Machine Stud.*, **7**, 511-528.

Ehrenstein, W. (1941) Über Abwandlungen der L. Hermannschen Helligkeitserscheinung. *Zeitschrift für Psychologie*, **150**, 83-91.

Ehrenstein, W. (1947) *Probleme der ganzheitspsychologischen Wahrnehmungslehre 2.* Auflage (Leipzig: Barth).

Ejima, Y., Redies, C., Takahashi, S., and Akita, M. (1984) The neon color effect in the Ehrenstein pattern. Dependence on wavelength and illuminance. *Vision Research*, **24**, 1719-1726.

Field, D., Hayes, A., Hess, R. (1993) Contour integration by the human visual system: Evidence for a local "association field." *Vision Research*, **33**, 173-193.

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

Gilchrist, A.L., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., Spehar, B., & Szura, J. (unpublished manuscript) A new theory of lightness perception.

Gregory, R.L., and Harris, J.P. (1974) Illusory contours and stereo depth. *Perception & Psychophysics*, **15**, 411-416.

Grossberg, S. (1994) 3D vision and figure-ground separation by visual cortex. *Perception and Psychophysics*, **55**, 48-120.

Grossberg, S., & Mingolla, E. (1985) Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, **92**, 173-211.

Guzman, A. (1968) Decomposition of a visual scene into three-dimensional bodies. In A. Grasselli (Ed.) *Automatic interpretation and classification of images*. New York: Academic Press.

Heider, G.M. (1933) New studies in transparency, form and color. *Psychologische Forschung*, **17**, 13-56.

Huffman, D.A. (1971) Impossible objects as nonsense sentences. *Machine Intelligence*, **6**, 295-323.

Jones, J., and Malik, J. (1992) A computational framework for determining stereo correspondence from a set of linear spatial filters. Proceedings ECCV, Genova.

Kanizsa, G. (1979) *Organization in vision*. New York: Praeger.

Kellman, P.J., & Shipley, T.F. (1991) A theory of visual interpolation in object perception. *Cognitive Psychology*, **23**, 141-221.

Kersten, D. (1991) Transparency and the cooperative computation of scene attributes. In Landy, M.S. & Movshon, J.A. (Eds.), Computational models of visual processing, 209-228. cambridge, Mass.: MIT Press.

Koffka, K. (1935) *Principles of Gestalt Psychology*. Harcourt, Brace, and World, Cleveland.

Kumar, T. (1995) Stereopsis due to luminance differences in the two eyes. *Vision Research*, **35**, 255-262.

Lawson, R.B., & Gulick, W.L. (1967) Stereopsis and anomalous contour. *Vision Research*, **7**, 271-297.

Lu, Z-L. & Sperling, George. (1996). 2nd-order illusions: Mach bands, Chevreul, and Craik-O-Brien-Cornsweet. *Vision Research*, **36**, 559-572.

Marr, D. and Poggio, T. (1976). Cooperative computation of stereo disparity. *Science*, **194**, 283-287.

Marr, D. and Poggio, T. (1979) A computational theory of human stereo vision. *Proceedings of the Royal Society of London (B)*, **204**, 301-328.

Mackworth, A.K. (1976) Model-driven interpretation in intelligent vision systems. *Perception*, **5**, 349-370.

McCourt, M.E. (1982) Spatial-frequency dependent brightness induction effect. *Vision Research*, **22**, 119-134.

McCourt, M.E. & Blakeslee, B. (1994) Contrast-matching analysis of grating induction and suprathreshold contrast perception. *Journal of the optical Society of America A*, **11**, 14-24.

Metelli, F. (1970) An algebraic development of the theory of perceptual transparency. Ergonomics, **13**, 59-66.

Metelli, F. (1974a) Achromatic color conditions in the perception of transparency. *In Perception: Essays in Honor of J.J. Gibson*, R. B. MacLeod, H. L. Pick (Eds.). Ithaca, NY: Cornell University Press.

Metelli, F. (1974b) The perception of Transparency. *Scientific American*, **230**, 90-98.

Metelli, F., da Pos, O., & Cavedon, A. (1985) Balanced and unbalance, complete and partial transparency. *Perception & Psychophysics*, **38**, 354-366.

Morinaga, S., Noguchi, K., & Osihi, A. (1962) Dominance of main direction in the apparent transparency. *Japanese Psychological Research*, **4**, 113-128.

Munker, H. (1970) *Farbige Gitter, Abbildung auf der Netzhaut und übertragungstheoretische Beschreibung der Farbwahrnehmung.* Habilitationschrift, Ludwig-

Maximilians-Universität München.

Nakayama, K., Shimojo, S. & Ramachandran, V.S. (1990) Transparency: relation to depth, subjective contours, luminance, and neon color spreading. *Perception*, **19**, 497-513.

Nakayama, K. and Shimojo, S. (1992) Experiencing and perceiving visual surfaces. *Science*, **257**, 1357-1363.

Pessoa, L. (1996) Mach bands: How many models are possible? Recent experimental findings and modeling attempts. *Vision Research*, **36**, 3205-3227.

Pessoa, L., Ross, W.D. (1996) A contrast/filling-in model of 3-D lightness perception: Benary cross, White's effect, and coplanarity. *Investigative Ophthalmology & Visual Science Supplement*, 4898.

Petry, S. & Meyer, G. (1987) The perception of illusory contours. New York: Springer-Verlag.

Petter, G. (1960) Osservazioni sulla trasparenza fenomenica. *Rivista di Psicologia*, **54**, 183-186.

Pollard, S.B., Mayhew, J.E.W., & Frisby, J.P. (1985) A stereo correspondence algorithm using a disparity gradient limit. *Perception*, **14**, 449-470.

Redies, C., Spillmann, L. (1981) The neon color effect in the Ehrenstein illusion. *Perception*, **10**, 667-681.

Redies, C., Spillmann, L., Kunz, K. (1984) Colored neon flanks and line gap enhancement. *Vision Research*, **24**, 1301-1309.

Rubin, E. (1921) *Visuell wahrgenommene Figuren*. Copenhagen: Gyldendals.

Schumann, M. (1904) Einige Beobachtungen Uber die Zusammenfassung von Gesichtseindruckern zu Binheiten. *Psychologische Studien*, **1**, 1-32.

Solomon, J.A., Sperling, G. and Chubb, C. (1993). The lateral inhibition of perceived contrast is indifferent to on-center/off-center segregation, but specific to orientation. *Vision Research,* **33**, 2671-2683.

Spehar, B. Gilchrist, A., and Arend, L. (1995) The critical role of relative luminance relations in White's effect and grating induction. *Vision Research*, **35**, 2603-2614.

Sperling, G. (1970) Binocular vision: A physiological and neural theory. *American Journal of Psychology*, **83**, 461-534.

Taya, R., Ehrenstein, W.H., Cavonius, C.R. (1995) Varying the strength of the Munker-White effect by stereoscopic viewing. *Perception*, **24**, 685-694.

Tuijl, H. F. J. M van (1975) A new visual illusion: neonlike color spreading and complementary color induction between subjective contours. *Acta Psychologica*, **39**, 441-445.

Tuijl, H. F. J. M van, Leeuwenberg, E.L.J. (1979) Neon color spreading and structural information measures. *Perception & Psychophysics*, **25**, 269-284.

Tuijl, H. F. J. M van, Weert C.M.M. de (1979) Sensory conditions for the occurrence of the neon spreading illusion. *Perception*, **8**, 211-215.

Varin, D. (1971) Renomeni di contrasto e diffusione cromatica nell'organizzazione spaziale del compo percettivo, *Rivista di Psicologia*, **65**, 101-128.

von der Heydt, R., Peterhans, E., Baumgartner, G. (1984) Illusory contours and cortical neuron responses. *Science*, **224**, 1260-1261.

von der Heydt, R., Peterhans, E., Baumgartner, G. (1989) Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *Journal of Neuroscience*, **9**, 1731-1748.

Ware, C. (1980) Coloured illusory triangles due to assimilation. *Perception*, **9**, 103-107.

Watanabe , T. & Cavanagh, P. (1993). Transparent surfaces defined by implicit X junctions. *Vision Research*, **33**, 2339-2346.

Watanabe, T., Takeichi, H. (1990) The relation between colour spreading and illusory contours. *Perception & Psychophysics*, **47**, 457-467.

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 gray bars within square-wave test gratings. *Perception*, **10**, 215-230.

# Footnotes

1. In textured images, it may not be the case that all of the luminance values within a shadowed region would be darker than all of the luminance values outside of the shadow. The constraint that shadows are darker would therefore only apply to the mean luminance of the texture. In this paper, we will only consider untextured images, so this subtlety is of no importance for the patterns considered herein.

2. Although there may be mechanisms sensitive to pure chromatic contrast that modulate the strength of the illusions described herein, the majority of the effects seem to be primarily (or at least strongly) shaped by the achromatic contrast relationships in the patterns.

3. To provide a greater appreciation for how the thesis of scission can be maintained without requiring that observers be explicitly aware of multiple surfaces, consider what occurs when observers attempt to make a lightness judgment. A lightness judgment --i.e., the recovery of surface pigmentation -- requires that observers have some means of determining the reflectance of surface and the prevailing illumination. Nonetheless, observers may have little conscious awareness that their judgments of surface lightness also expresses a judgment about the illuminant. It is in this sense that the thesis of scission need not entail that observers be aware of the decomposition of a region into multiple contributions.

4. There have been some reports of perceived transparency in displays that contain double reversing X-junctions (Beck et al., 1994). We will discuss this issue in Section 5 of this paper.

5. However, some of these issues will be discussed in Section 4 when discussing how illusory contours are generated in binocular displays, because no existing theory is capable of explaining how illusory contours are generated in the binocular displays that we will consider in this section.

6. The paradox arises due to the failure of contrast enhancing mechanisms -- such as lateral inhibition -- in predicting the correct sign of this illusion.

7. Perhaps the simplest way to understand why an assumption of this kind might be implemented by the visual system is that such a decomposition would minimize the number of contrast changes used to explain the data, and in this sense, may be considered to be the "simplest" form of scission.

8. In a previous version of this paper, a reviewer pointed out that this line of argument does not provide conclusive evidence that the Munker-White illusion involves scission. As noted by the reviewer, even if stereoscopic depth differences are shown to enhance the Munker-White illusion in the direction predicted by scission, this does not necessarily imply that scission is the correct explanation of monocular variants of this illusion, since some other mechanism may be responsible for the monocular effects. I am in full agreement with this line of reasoning. However, at this juncture, scission is the only theory offered to date that provides a sufficient account of the relative magnitude (compared with simultaneous contrast) and direction of the monocular illusion. Therefore, the stereoscopic data should simply be viewed as contributing support for the scission thesis.

9. At least one reviewer was concerned with the possibility that observers were in some way "led" to this observation by the experimenter. In fact, just the opposite occurred. The experimenter (B.A.) was incapable of understanding how an experienced observer in a neighboring lab could report the "behind" configuration as stronger than the monocular form of this illusion, since the perceptual organization of the display was presumably now equivalent to a simultaneous contrast display. When questioned about this possibility, the observer (Y.W.) commented that he did not experience the display as identical to a simultaneous contrast display occluded by a series of bars, but rather, reported that some of the luminance of the stripes in which the targets were embedded appeared as faint transparent filters. I then questioned the other anomalous

observer whether any portion of the display generated percepts of transparency, and he (S.G.) reported a similar percept. Therefore, any "demand characteristics" that were introduced were inflicted by the observers on the experimenter, not vice-versa. I have subsequently tested five naive observers, but none reported the enhancement of the illusion when the targets were behind the adjacent bars, so I have not be able to provide additional evidence to support this interpretation.

10. It may be argued that this figure also induces monocular percepts of illusory contours, and that the binocular illusory contours are caused by monocular illusory contours that input to conventional binocular disparity mechanisms. However, we have also performed discrimination experiments the reveal that illusory contours can be formed in displays that contain a single binocular terminator generated by a polarity reversing contour, and that no monocularly visible illusory contours are present in these displays (cf. Anderson, 1994; Anderson & Julesz, 1995).

11. It should be noted that this assertion depends critically on how the discrepant results reported by Bressan (1993) and van Tuijl & de Weert (1979) are reconciled. However, from a historical perspective, it is worth noting that the Grossberg & Mingolla model was inconsistent with the data existing at the time of the model's formation (van Tuijl & de Weert, 1979). Depending on how the discrepant experimental reports are resolved, this will either prove to be a critical flaw or a predictive triumph of their model.