Perception, 1997, volume 26, pages 471 – 492

Color transparency

Michael D'Zmura¶, Philippe Colantoni, Kenneth Knoblauch, Bernard Laget
Institut d'Ingénierie de la Vision, Université de St Etienne, St Etienne, France
Received 10 June 1996, in revised form 24 February 1997

Abstract. Observation suggests that the chromatic changes which elicit an impression of transparency include translations and convergences in color space. Neither rotations nor shears in color space lead to perceived transparency. Results of matching experiments show that equiluminous translations, which cannot be generated by episcotister or filter models, give rise to the perception of transparency. This implies that systematic luminance change is not needed for transparency to be perceived. These results were used for the development of a method for detecting a transparent overlay within a color image and for separating the overlay from the underlying surfaces. The method tests for the coherence of chromatic change along contours through X-junctions to help detect the contour of a transparent region. The algorithm tests locally for translation and convergence to detect a transparent region. It estimates globally the chromatic parameters of the transparent overlay in order to separate the overlay from the underlying surfaces.

1 Introduction

The visual system identifies surfaces and uses their properties to help recognize objects. Surface color is an important source of information, despite the frequent changes in the chromatic properties of lights from surfaces. For example, shadows on a sunny day cause lights from surfaces to become darker and bluer. Surface specularity causes incident lights to be mirrored back towards the viewer as highlights, in addition to the diffusely reflected light that bears the surface color (Shafer 1985; D'Zmura and Lennie 1986). A change in illumination across a scene causes reflected lights to change in a way that depends on the chromatic properties of the illuminants; the proper interpretation of this change underlies color constancy (D'Zmura 1992).

The visual system uses surface color in identification and recognition despite such chromatic variation because it readily interprets systematic chromatic change correctly as a change in surface viewing conditions (Craven and Foster 1992; Foster et al 1992; D'Zmura and Mangalick 1994). Color transparency provides an excellent example of this (see figure 1a). A surface that is seen both in plain view and through a transparent overlay is identified as a single surface, despite the variation in its lights. Observers separate the color properties of a transparent filter from the color properties of underlying surfaces.

What conditions on colors in a still image are needed to perceive transparency? The first is that two or more surfaces are seen both in plain view and under a transparent overlay (Tudor-Hart 1928). The border between two such surfaces will create a point of intersection between four different colored regions as it crosses under the overlay. The second condition is that the changes in color at these four-way intersections or X-junctions must be such as to lead the visual system to infer that surface regions seen in plain view correspond to surface regions seen under the overlay: systematic color change is required for the visual system to infer surface correspondence.

The systematic color changes needed for the perception of transparency have been identified as additive and subtractive color mixture in previous work (Metelli 1974;

¶ Address for correspondence: Michael D'Zmura, Department of Cognitive Sciences, University of California, Irvine, Irvine, CA 92697, USA; e-mail: mdzmura@uci.edu.

Beck 1978; Beck et al 1984; Da Pos 1989). Metelli (1974) used achromatic images to show that additive mixtures of lights can lead to strong impressions of transparency. These additive mixtures generalize Talbot's law of color fusion, which describes the light mixtures met when using an episcotister apparatus (Beck 1978). In this form of transparency, lights from surfaces are mixed additively with the light from a transparent overlay to produce the lights from the surfaces when viewed through the overlay. Additive color mixture plays a role in both perception and performance. Da Pos (1989) studied the perception of colored stimuli to generalize Metelli's results with achromatic additive color mixture to the three dimensions of color space. De Weert (1986) tested the ability of observers to read overlapping pairs of words and found that words are most legible when their values are mixed additively.

Beck and his colleagues argued for the importance of subtractive color mixture (Beck 1978; Beck et al 1984). They pointed out that a light-absorbing filter which is situated in front of a set of surfaces absorbs or subtracts light (Beck 1978). Indeed, the spectral properties of lights from surfaces that lie under such a filter can most easily be modelled as products of surface and filter spectral properties; light that is directly reflected from the filter towards the viewer can also contribute an additive component (Brill 1984; Kersten 1991; Mulligan 1993).

Gerbino and colleagues (1990) used a model of an episcotister to develop formulae for transparency that combine both additive and subtractive color mixture. Their model unifies much earlier work on conditions for achromatic transparency perception. The work that we present here extends the unified model to color transparency.

We borrow from research on optic flow (eg Gibson 1979; Koenderink 1986) to characterize the changes in the chromatic properties of lights which lead most effectively to the perception of transparency. One may think of systematic changes in three-dimensional color space in terms of vector fields that describe motions. The base of each vector represents the chromatic properties of a surface seen in plain view, while the arrowhead represents the chromatic properties of the same surface seen under the transparent overlay. Many such surfaces provide a field of vectors that may represent systematic motions, including translation, convergence or divergence, rotation (curl), and shear (deformation). Which of these systematic motions in color space leads to the perception of transparency?

We report here that translation, convergence, and composition of the two can give rise to the perception of transparency. Translation and convergence correspond to the cases of subtractive and additive color mixture (Beck 1978; Beck et al 1984), respectively. By combining translation and convergence in color space, we generalize the model of Gerbino et al (1990) to color. We have found that other systematic changes in color space, including rotation and shear, do not lead to the perception of transparency.

We use these results to create an algorithm for detecting transparent overlays in color images. It tests locally whether X-junctions along a candidate border of a transparent region are consistent with translation and convergence. Performing local tests of this sort lets the algorithm detect both spatially uniform transparent overlays and overlays with space-varying chromatic properties. The algorithm also estimates parameters of the global chromatic change across the border of a transparent region. If the overlay is homogeneous, then the algorithm uses these parameters to separate overlay chromatic properties from those of the underlying surfaces.

2 Chromatic change

We examine whether translation, convergence, rotation, and shear (Schey 1992) in color space lead to an impression of transparency.

2.1 Translation

Figure 1a (see color plate 1 on page 475) shows a flat array of Color Aid® papers over which has been laid a small square sheet of transparent yellow plastic. Seven of the papers lie on the border of the yellow filter and so are seen both in plain view and through the filter. Figure 1b contains plots of the changes in the chromaticities of the lights from these seven papers. Each vector represents one of the seven papers: its base marks the chromaticity of the paper when seen in plain view and its arrowhead marks the chromaticity when seen through the filter.

The plot shows that the changes in chromaticity are larger for some surfaces (eg the blue ones), and smaller for others (eg the red) and that the directions of change differ but are quite similar. The simplest model for the dependence of such changes on surface and filter properties is a multiplicative one: functions of wavelength that describe surface and filter spectral properties, respectively, are multiplied together to produce the spectral properties of the light that reaches the viewer (eg Brill 1984; Kersten 1991). Such models have the same bilinear form as models used in the study of color constancy (eg Maloney and Wandell 1986; D'Zmura and Lennie 1986). While analysis of such models shows that precise details of the change in chromaticities of surfaces can be informative (D'Zmura 1992; D'Zmura and Iverson 1997), we avoid the complexities of these models here. A translation in color space is a good, first-order approximation to the shifts in lights from surfaces when viewed under a filter.

The plot in figure 1b shows that the papers shift chromaticity in approximately the same way, each becoming more yellow. Not shown are the accompanying decreases in luminance of each paper. Indeed, the light-absorbing yellow filter causes each paper to become more dark yellow. The overall change can be approximated well by a rigid shift—namely a translation in color space.

A light-absorbing filter causes less light to reach the eye from an underlying surface. One might guess that a translation in color space, like the change in figure 1, must include a decrease in luminance if transparency is to be perceived. Yet cases in which luminance is increased can also lead to a strong impression of transparency (Metelli 1974). One might argue that an area of increased luminance is actually an area in which surfaces are seen in plain view, and that a filter with a hole cut out covers the rest of the display. In this view, there is still a luminance decrease, caused by a light-absorbing filter in the surround, that is responsible for the perceived transparency.

This fails to explain the perception of transparency with equiluminous translations. We produced examples of chromatic translations along the equiluminous L&M-cone and the S-cone axes of color space (MacLeod and Boynton 1979; Krauskopf et al 1982; Derrington et al 1984). We used displays of simulated surfaces which had reflectance functions that were generated by a stochastic linear model described in Appendix A. The surfaces were lit, in simulation, by CIE daylight illuminant D_{65} (Wyszecki and Stiles 1982). The computed lights were then presented on a computer monitor that was calibrated and gamma-corrected.

The chromatic translations were applied to a central square region in the simulated images. We first shifted the chromaticities of reflected lights towards the 'red' end of the L&M-cone axis at equiluminance, or towards the 'blue-green' end of the same axis. Both translations lead one to perceive transparency, despite the absence of systematic luminance changes in the lights from the surfaces. The central square area in figure 2a (see color plate on page 475) shows an example of a chromatic translation towards the 'red' end of the L&M-cone axis. Figure 2b shows the uniform shifts in chromaticity of the lights from the surfaces along the border of the transparent area, which are seen both in plain view and through the filter. We hasten to point out that this and other color figures are unlikely to reproduce exactly the calibrated images. We also created displays with a central square region shifted in chromaticity along the S-cone axis towards

either the 'yellow-green' half of the axis or the 'purple' half. These equiluminous shifts, too, lead one to perceive transparency.

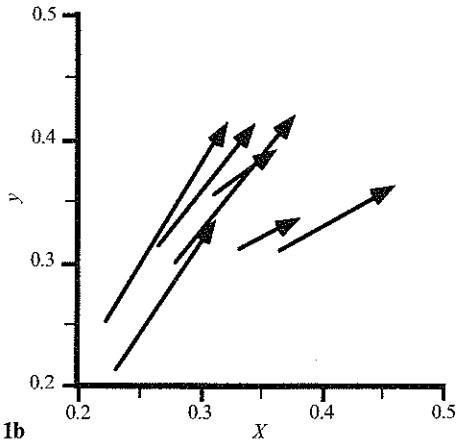


Figure 1. Color transparency. (a) (color plate 1, facing page) Image of a yellow filter lying atop an array of Color Aid® papers, acquired by an RGB CCD camera; (b) shifts in chromaticities of the lights from the papers in the CIE 1931 (x, y) chromaticity plane. The base of each arrow represents the chromaticity of a surface seen in plain view; the head of the arrow represents the chromaticity of the same surface seen through the yellow filter.

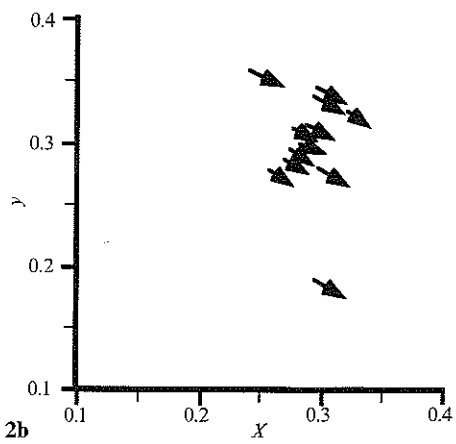


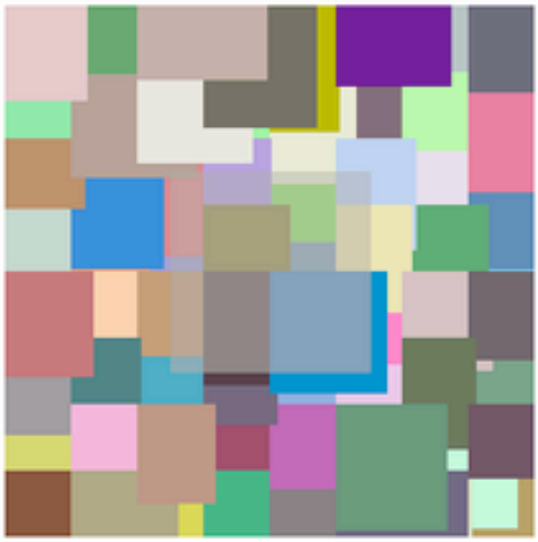
Figure 2. Transparency with an equiluminous translation. (a) (color plate 1, facing page) Image of a translation towards the 'red' end of the L&M-cone axis. An equiluminous modulation of color that is invisible to the short-wavelength-sensitive S-cones is applied to lights from a central square region. The modulation, applied to a neutral gray surface of average intensity, provides a contrast of about 6% to L-cones. (b) Shifts in CIE 1931 (x, y) chromaticities of the lights from the simulated surfaces.

2.1.1 *Transparency matches for equiluminous translations*

We conducted experiments to measure the perceived strength of transparency caused by shifts in chromaticity at equiluminance. Observers matched the strength of perceived transparency of an equiluminous shift to that of an achromatic standard.

2.1.1.1 *Methods.* Stimuli were presented on an Eizo FlexScan T562-T color monitor, which observers viewed binocularly at a distance of 0.56 m in a dark room. Software on a Compaq Prolinea 486 controlled a Cambridge Research Systems VSG/2 color graphics board which was set to provide 12 bits of chromatic information for each of the 800×600 pixels presented on the monitor at a field rate of 100 Hz (noninterlaced). The display subtended 32 deg wide by 24 deg high. The nonlinear relationship between applied voltage and phosphor intensity was corrected, for each gun, with the aid of color lookup tables. The chromaticities and luminances of the three phosphors of the monitor were measured with a Minolta CS-100 chromameter. The mean luminance of the screen was 65 cd m^{-2}; its mean chromaticity was (0.294, 0.303) for the CIE 1931 standard observer.

On the screen was displayed an 8 wide by 6 high array of colored squares (see figure 3). The R, G, and B values of the squares were chosen randomly by independent draws from a uniform distribution that permitted a maximum excursion of 50% contrast. On a scale of 0–100, R, G, and B values each varied, independently, in the range 25–75; the average gray of the display had R, G, and B values of 50, 50, and 50. Superimposed on the array were two square regions in which the chromatic properties of underlying squares were translated in color space. One of these superimposed squares was generated by decreasing the luminosity of underlying pixels, and this square served as the standard achromatic example of transparency. The position of this achromatic standard, at either left or right, was varied randomly from trial to trial.



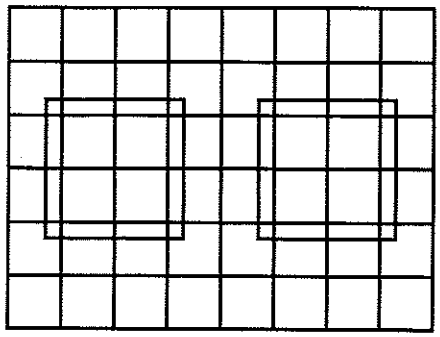


Figure 3. Spatial properties of the display used in the experiments on transparency matching. Each square in the 8 wide by 6 high array subtended 4 deg × 4 deg of visual angle. The two larger squares in which chromaticities were translated are of size 10 deg × 10 deg.

The other superimposed square was generated by an equiluminous color shift, and observers varied the chromatic contrast of this shift in an attempt to match the strength of its perceived transparency to that of the achromatic standard.

Three color-normal male observers participated in the experiments. For each of four hues, ten trials at each of six achromatic contrast levels (0.05, 0.075, 0.1, 0.125, 0.15, and 0.175) were presented in random order to the observers. Observers adjusted the physical color contrast of the equiluminous translation so that the two transparent squares provided equally strong impressions of transparency. Observers pushed a switch to signal that they had found a satisfactory match. At the highest two achromatic contrast levels, 0.15 and 0.175, it was often impossible for observers to make a match within the available range of chromatic contrast. In such cases, observers pushed another switch to signal that no match was possible.

The four hues chosen were the red and the blue-green of the L&M-cone axis, modulations along which are equiluminous and visible only to L and M cones sensitive to the long and medium wavelengths, and the yellow-green and purple of the S-cone axis, modulations along which are equiluminous and visible only to the short-wavelength-sensitive S cones (Smith and Pokorny 1975; Derrington et al 1984).

2.1.1.2 *Results.* The areas of achromatic translation looked like neutral density filters, while the areas of equiluminous translation looked like colored but otherwise clear filters. Matches were easy to set. Figure 4 shows the color contrasts that are needed for equiluminous translations to produce impressions of transparency that are equal in perceived strength to those caused by achromatic decrements. Results are shown for three observers (MD, top; PC, middle; and VL, bottom). Results for red and blue-green shifts along the L&M-cone axis are shown by open and filled circles, respectively, at the left. Results for yellow-green and purple translations along the S-cone axis are shown by open and filled circles, respectively, at the right. The contrasts are shown in units appropriate to the chromatic axis and are measured relative to the average gray level of the display. Contrast along the achromatic axis (bottom) is in standard units, while contrast along the L&M-cone axis (left) is in units of contrast to the L-cones, and contrast along the S-cone axis (right) is in units of contrast to the S-cones.

For each observer and each hue, matches increase monotonically with increasing contrast of the achromatic standard. Along the L&M-cone axis (left), as the contrast of the achromatic standard was increased from 0.05 to 0.125, observers MD and VL were able to set a match using color contrasts in the available range. These observers could not match the achromatic standard at higher contrasts because of color gamut limitations. Observer PC could make matches within the range 0.05 to 0.1, but above this range of achromatic contrasts matches were impossible to set. Similar results were found by the three observers when using equiluminous translations along the S-cone axis (right). With one exception, the amount of chromatic contrast needed to match was nearly the same along opposite directions of a chromatic axis: the open

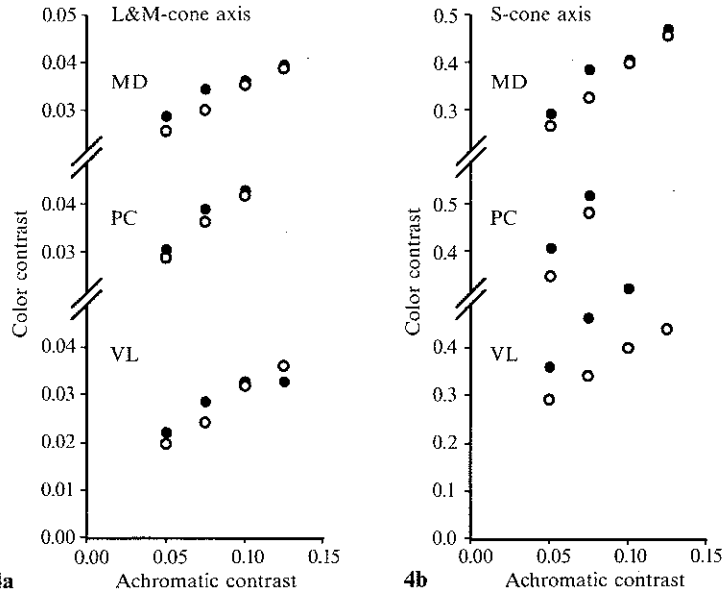


Figure 4. Results of experiments in which the perceived strength of transparency that arises from translations at equiluminance was matched to the perceived transparency of achromatic standards. Three observers—MD, PC, and VL—matched equiluminous translations to luminance decrements presenting contrasts of 0.05 through 0.15 (horizontal axes). (a) L&M-cone axis matches, both for the 'red' half of the axis (open circles) and the 'blue-green' half of the axis (filled circles). Units of contrast along the vertical axis are in terms of contrast to L-cones. (b) S-cone axis matches, both for the 'yellow-green' half of the axis (open circles) and the 'purple' half of the axis (filled circles). Units of contrast along the vertical axis are in terms of contrast to S-cones.

and filled circles lie atop one another. The exception is for observer VL when viewing S-cone axis stimuli. In this case, purple translations were less effective than yellow-green ones in providing transparency.

The ease with which the observers were able to make reliable matches of perceived strength of transparency supports our informal observations that equiluminous translations lead to the perception of transparency.

2.1.2 *Formulation of translation*

We formulate translation in the CIE XYZ color space of human observer tristimulus values (Wyszecki and Stiles 1982; Da Pos 1989). With this color space, one can often obtain results that are useful in the many color spaces that are related by a linear transformation (cf Brainard 1995). For example, the *RGB* color spaces provided by the phosphors of most television sets and computer monitors are related linearly to XYZ; so, too, is the color space *LMS* of human cone fundamentals (Smith and Pokorny 1975).

Suppose, then, that a surface reflects a light that is represented by a vector of tristimulus values $\boldsymbol{a} = (X_a\, Y_a\, Z_a)$. A translation is a vector of tristimulus values $\boldsymbol{t} = (X_t\, Y_t\, Z_t)$ that is added to the tristimulus values of the surface $\boldsymbol{a}$ to provide new tristimulus values $\boldsymbol{b} = (X_b\, Y_b\, Z_b)$:

$$\boldsymbol{b} = \boldsymbol{a} + \boldsymbol{t}. \tag{1}$$

Identical translations of the lights from many surfaces provide a systematic change in color that is readily interpreted by the visual system. Our observations with equiluminous translations and the results of matching experiments show that systematic changes in luminance across the border of a connected region are not necessary for the perception of transparency. Translation in any direction of color space will work.

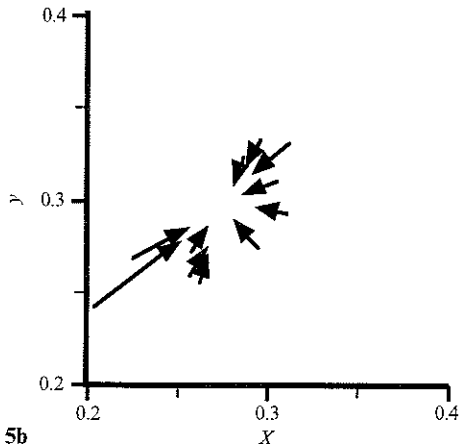
2.2 Convergence

The work of Metelli (1974) with achromatic stimuli and that of Da Pos (1989) with colored stimuli show that convergence in color leads to robust perceptions of transparency. Such convergence is shown clearly by achromatic stimuli wherein lighter surfaces become darker and darker surfaces become lighter: surface lightnesses in the transparent region converge towards a central gray (eg figure on page 43 of Metelli 1974; figure 11 of Da Pos 1989). The region of convergence is perceived not only as overlaid by a transparent surface but also as having lesser contrast than the surrounding area. This holds true also for heterochromatic examples of convergence (eg plates 3–6 of Da Pos 1989). We show an example of convergence towards a neutral gray point in figure 5. The chromatic changes undergone by the surfaces represented in figure 5a (see color plate 1 on page 475) are shown in figure 5b, and the motion towards the center of convergence is clear.

The point towards which colors converge can vary over a wide range and still lead to the perception of transparency. We generated displays of convergence, like that in figure 5, for centers of convergence in six regions of color space, and present examples in figures 6–8 (color plate 2, facing page). Figure 6a shows convergence towards 'dark gray', while figure 6b shows convergence towards 'light gray'. Figure 7a shows convergence towards 'red' and figure 7b towards 'blue-green' along the L&M-cone axis. Figure 8a shows convergence towards 'yellow-green' and figure 8b towards 'purple' along the S-cone axis. The perception of transparency is robust in all cases.

The opposite of a convergent motion is a divergent one. Informal observation shows that divergence also leads to an impression of transparency, yet for many this transparency is more strongly associated with the (convergent) background in which contrast has been decreased rather than with the divergent figure.

It is of interest to determine whether systematic luminance changes are needed for convergence to produce an impression of transparency. We examined this informally using arrays of equiluminous rectangles. Within a central region, the color contrasts of the rectangles were reduced to produce an equiluminous convergence of color. One may still perceive transparency under these wholly equiluminous conditions, which are



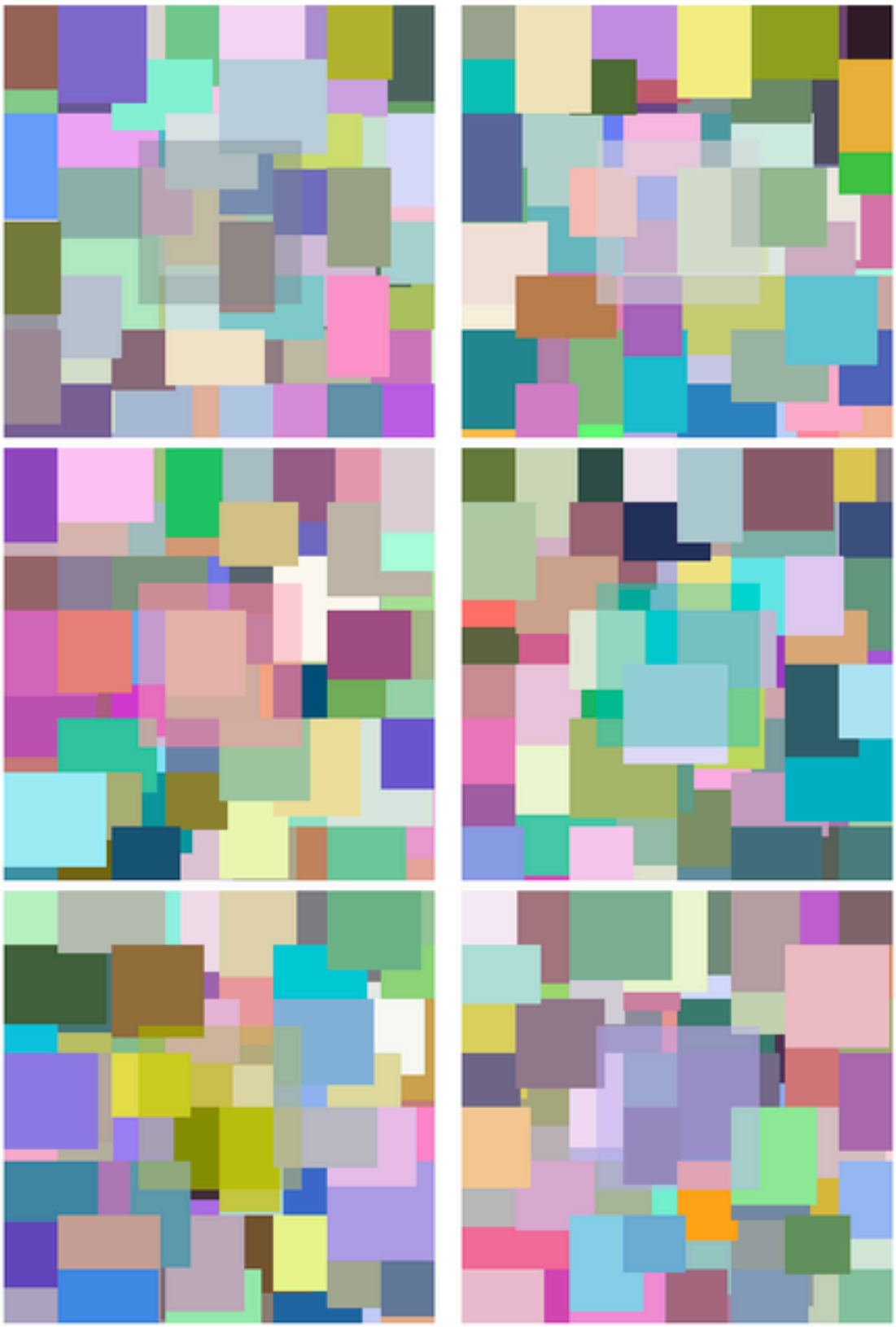
5b

Figure 5. Convergence towards a neutral point in color space. (a) (color plate 1, page 475) Color image. (b) Plot of chromaticity shifts in lights from simulated surfaces along the border of the central square area.

Figure 6. Convergence towards achromatic targets (color plate 2, facing page). (a) Convergence towards a dark gray. (b) Convergence towards a light gray.

Figure 7. Convergence towards targets along the L&M-cone axis (color plate 2, facing page). (a) Convergence towards red. (b) Convergence towards blue-green.

Figure 8. Convergence towards targets along the S-cone axis (color plate 2, facing page). (a) Convergence towards yellow-green. (b) Convergence towards purple.



similar to ones used in studies of contrast induction (Singer and D'Zmura 1994, 1995; D'Zmura and Singer 1996).

2.2.1 *Formulation of convergence*

We follow Metelli (1974) and Da Pos (1989) in formalizing chromatic convergence as movements of surface colors towards some target color along the line segments in color space joining the surface colors to that of the target. Suppose that a surface color is represented by a vector $\boldsymbol{a}$ of tristimulus values and that the target of convergence is represented by vector $\boldsymbol{g}$. Then the result is a vector $\boldsymbol{b}$ of tristimulus values that lies along the line segment that joints $\boldsymbol{a}$ and $\boldsymbol{g}$. This result $\boldsymbol{b}$ can be written as a linear combination of $\boldsymbol{a}$ and $\boldsymbol{g}$ in the following parameterized form:

$$\boldsymbol{b} = (1-\alpha)\boldsymbol{a} + \alpha\boldsymbol{g}, \quad 0 \leqslant \alpha \leqslant 1. \tag{2}$$

The parameter α determines the extent to which the tristimulus values $\boldsymbol{b}$ are changed from $\boldsymbol{a}(\alpha = 0)$ to $\boldsymbol{g}(\alpha = 1)$ in convergent motion (or away from $\boldsymbol{g}$ in divergent motion with $\alpha < 0$). Convergent movement of the lights from many surfaces towards a single target provides a systematic change in color that is readily interpreted by the visual system.

Watanabe and Cavanagh (1993) note that it is possible for the lights from a surface, seen both in plain view and under a transparent overlay, to undergo no change whatsoever. This can occur in the case where the original surface tristimulus values $\boldsymbol{a}$ are identical to those of the target of convergence. Watanabe and Cavanagh point out that the T-junctions found with such surfaces function as 'implicit X-junctions' and that these can give rise to impressions of transparency.

2.2.2 *Compositions of translation and convergence*

It is possible to create chromatic motions that have both translatory and convergent elements. Such composite motions create both a uniform shift in the colors of surfaces and a reduction of contrast, just like a convergence to a non-neutral point, as in figures 6–8. Indeed, the result of composing a translation and a convergence in color space is a convergence. This can be shown as follows. Suppose that to a convergence, like that expressed in equation (2), is added a translation $\boldsymbol{t}$. The resulting color will then be represented by a vector $\boldsymbol{b}$ of tristimulus values that is given by the following formula:

$$\boldsymbol{b} = (1-\alpha)\boldsymbol{a} + \alpha\boldsymbol{g} + \boldsymbol{t}, \quad 0 \leqslant \alpha \leqslant 1. \tag{3}$$

Yet this can be rewritten as a convergence as follows:

$$\boldsymbol{b} = (1-\alpha)\boldsymbol{a} + \alpha\boldsymbol{g}', \tag{4a}$$

in which the new center of convergence $\boldsymbol{g}'$ is given by

$$\boldsymbol{g}' = \boldsymbol{g} + \frac{1}{\alpha}\boldsymbol{t}. \tag{4b}$$

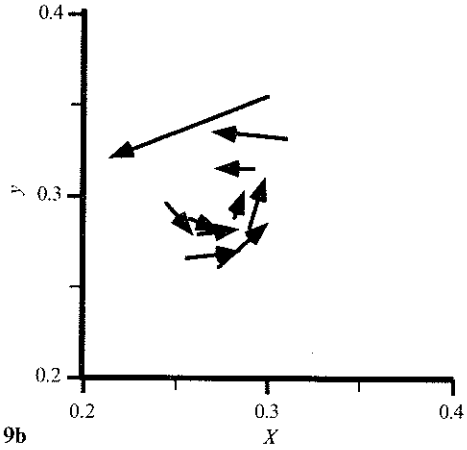
This result is not surprising if one recalls that a translation is a convergence to an infinitely distant point. A similar calculation shows that the composition of two convergences is yet another convergence, from which one concludes that the set of convergences is closed under composition.

2.3 Curl and shear

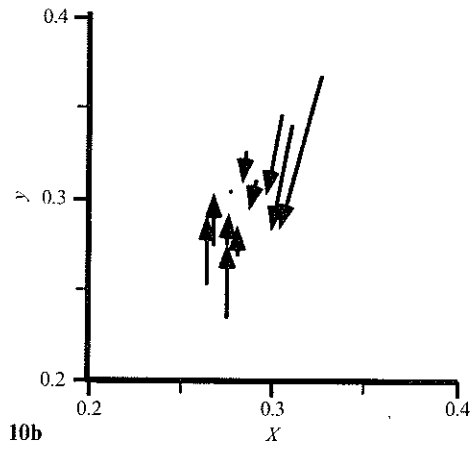
Translation and convergence are special motions in consequence of the ability to express them with the sole use of colors that lie along a single axis in color space. This axial nature of the motions is evident if one considers translation and convergence in achromatic images, in which all colors lie along the achromatic axis.

Two other fundamental motions in three-dimensional vector spaces include rotation (curl) and shear (deformation). Neither rotation nor shear, however, can be expressed as motions among colors that lie along a single axis and, as we shall see, neither leads to robust perception of transparency.

Figure 9 shows an example of chromatic rotation. The image in figure 9a (see color plate 1, page 475) shows surfaces with chromaticities that rotate about a neutral gray as they pass from plain view to a position under the region of rotation. Visible are a blue surface that becomes red (within the putative region of transparency), a red surface that becomes orange, an orange surface that becomes yellow-green, and a yellow-green surface that becomes green. The corresponding chromaticity changes are shown in figure 9b. It is possible to use smaller or greater rotations in color space to create images with less rotation or more rotation. But in all cases of rotation that we have investigated the transparency percept is broken. There may be areas within the putative overlay figure in which transparency is perceived, but these areas are formed by the fortuitous adjacency of surfaces with colors that change in the same direction: the transparency that is perceived can be attributed to accidental translation.



9b



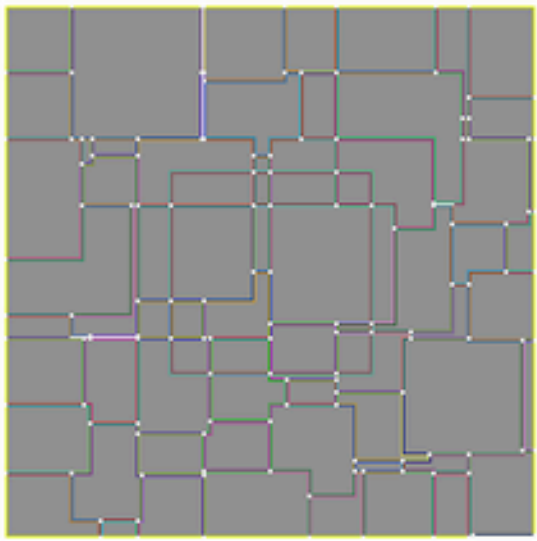
10b

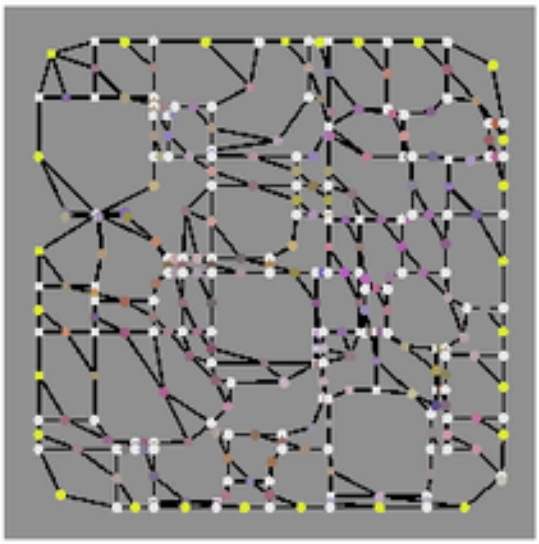
Figure 9. Chromatic curl. (a) (color plate 1) Color image. (b) Change in chromaticities. See text for discussion.

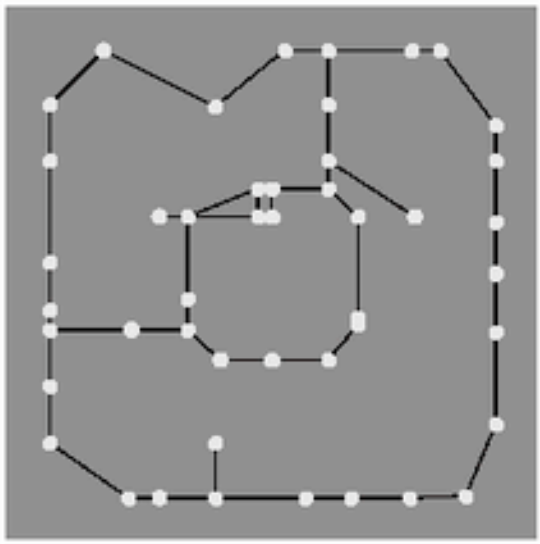
Figure 10. Chromatic shear. (a) (color plate 1) Color image. (b) Change in chromaticities. See text for discussion.

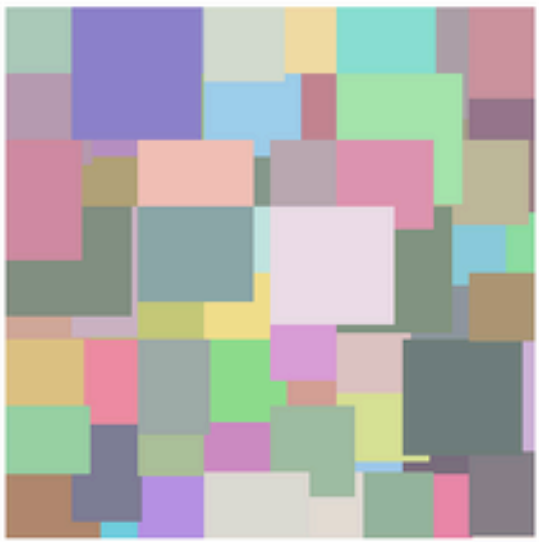
Figure 10 shows an example of chromatic shear. The image in figure 10a (see color plate 1, page 475) shows surfaces with chromaticities that are deformed about a neutral gray as they pass under the putative region of transparency. In this example, yellowish surfaces become more red, and bluish surfaces become more green. Surfaces with colors along a red-green axis undergo no change. Figure 10b shows the change in chromaticities. As was the case with rotation, the transparency percept is broken. In all examples of shear that we have investigated, one can attribute the transparency that remains to an accidental adjacency of surfaces that leads to transparency through translation.

Please note that, because our observations have been less than exhaustive, we cannot rule out the possibility that there may be certain rotations and shears that lead directly to the perception of transparency. However, such examples remain to be created.









3 Detection and separation

The results of the observations suggest that we can detect transparent overlays by examining X-junctions with chromatic changes that are consistent with translation or convergence or both. Below, we present an algorithm that operates on color images to detect such transparent regions and then separate the chromatic properties of the overlay from those of underlying surfaces. The steps in the algorithm are:

- identification of contours and intersections;
- creation of an X-junction adjacency graph (XJAG);
- determination of the contour of the transparent overlay by using spatial smoothness and chromatic coherence constraints;
- separation of the transparent overlay from underlying surfaces.

We consider now each of these in turn.

3.1 Identification of contours and intersections

The algorithm uses region growing to segment the image into bounded regions of homogeneous color (Gonzalez and Wintz 1977; Russ 1995). One grows a region around a starting pixel by including neighboring pixels that have a color sufficiently similar to that of the starting pixel. The region boundaries form contours. For each pixel along these contours, we determine the direction of the spatial tangent vector and the change in color across the contour. To identify intersections, we count the number of distinct colors in the neighborhood of each pixel. T-junctions have three neighboring colors and X-junctions have four.

Figure 11 (color plate 3, facing page) shows the result of applying this first stage of processing. The input image, shown in figure 11a, is a simulated image with a translation in a dark-red direction. The contour and intersections are shown in figure 11b. We represent each contour here by a pair of colors centered chromatically about neutral gray, the difference between which is identical to the chromatic difference across the contour. Each intersection is marked by a white spot.

Determining the positions and chromatic properties of region contours is critical to the algorithm, which depends on the analysis of X-junctions and their chromatic changes. Segmentation through region growing works well on simulated images of the sort in figure 11a: each region is homogeneous and the contours are sharp. The segmentation works far less well with regions that are nonuniform or that lack sharp edges. This problem is raised again in section 4.

3.2 Creation of an X-junction adjacency graph (XJAG)

We use the contour and intersection to create an X-junction adjacency graph (XJAG). Towards this end, we first create a contour adjacency graph (CAG), in which contours and junctions are represented by nodes, and adjacency among contours and intersections is represented by links among the nodes (see figure 12a, color plate 3, facing page). The nodes are represented by colored spots and the links by black line segments. Each white spot represents an intersection—either a T-junction or an X-junction. The yellow spots around the border of the image represent T-junctions along the border, which we

Figure 11. Algorithm: Input and contour detection (color plate 3, facing page). (a) Input image. (b) Color difference contour with intersection image. The spurious white line segments in the latter are caused by very thin regions in the input picture; the test for the number of neighboring colors for pixels within such regions has returned the value 3 (T-junction) for each pixel in several of these regions.

Figure 12. Algorithm: Adjacency graphs (color plate 3, facing page). (a) Picture of a contour adjacency graph (CAG) for the image in figure 11a. White spots represent intersections; yellow dots along the border represent T-junctions along the image border. Other colored spots represent contours. The links represent adjacency. (b) Picture of an X-junction adjacency graph (XJAG) for the image in figure 11a. White spots represent X-junctions, and links the adjacency relationships among the mediated by image contours.

include in our analysis in order to be able to detect transparent overlays that lie along the image border. The other spots represent contours, and the links in the graph represent spatial connection between these interactions and contours. Such a graph is similar in spirit to a region adjacency graph, in which regions determined by segmentation are represented by nodes, and adjacency between regions is represented by links (Pavlidis 1980; Schettini 1993). The difference is that the nodes in the CAG are region contours and intersections rather than the regions themselves.

The XJAG is formed from the CAG by (i) changing the representation of contours from nodes to links and then (ii) removing T-junctions, unless they are on the border of the image (see figure 12b, color plate 3, page 482)). Each node (white spot) represents either an X-junction or a T-junction along the image border. The links in the XJAG contain adjacency information, the spatial tangent vectors of represented contours, the chromatic change vectors, and contour labels.

There are several possible circuits through the X-junctions represented in the XJAG of figure 12b. One such closed circuit is the path through the T-junctions along the border, which includes the entire image. There are several circuits through the X-junctions in the center of the image; one of these circuits corresponds to the transparent overlay. In addition, there are several circuits which include the T-junctions in the upper-left corner of the image and pass through the center of the image and, finally, several circuits which include the T-junctions in the lower-right corner and pass through the center of the image. Even very simple images like the one in this example (figure 11a, color plate 3, page 482) typically contain many circuits.

3.3 Determination of contour through analysis of spatiochromatic change

The algorithm attempts to find, if possible, a closed circuit through junctions that lie in the center of the image that satisfy smooth contour and coherent chromatic change constraints. Failing this, it searches for circuits that include T-junctions along the image border and junctions in the image center that meet the constraints.

3.3.1 *Smooth continuation*

The algorithm uses the spatial tangent vectors to the contours in the XJAG to find the smoothest contour through the X-junctions. The smoothest continuation through an individual X-junction is determined by evaluating changes in contour direction on either side of the X-junction. If the smoothest circuit includes a change in contour direction that is too great (the difference in angle is larger than some threshold), then the circuit is classified as not smooth.

In many cases, the smooth-continuation constraint effectively identifies a smooth contour through central X-junctions that corresponds to the border of the transparent region. This is far from sufficient, however. Consider a square array of randomly colored regions. At each corner one would find an X-junction. The algorithm would find many smooth circuits through X-junctions in the center of the image, and most likely none of these circuits would correspond to a transparent region. One could color several of the squares in a way that leads to the impression of transparency, however. For the algorithm to detect this transparent region, it must evaluate chromatic change along each circuit, both to eliminate regions that are not transparent and to detect ones that are.

3.3.2 *Chromatic coherence*

There are many ways to test whether chromatic changes are consistent with translation or convergence in color space. A first distinction among methods is whether tests are applied locally and individually to the four colors found at each X-junction, or globally and simultaneously to all colors found along the candidate contour of a transparent region. We use both methods. The desirability of detecting transparent

overlays that have spatially heterogeneous chromatic properties prompts us to use local tests for detection. The need to estimate the chromatic parameters of a transparent overlay leads us to use global methods for identification. A second issue is precision. For example, how much of a deviation from a rigid shift in color space should the algorithm tolerate in a candidate translation? We propose rules that ignore the magnitudes of color changes and test only whether color changes are in the proper direction, within some angular tolerance.

3.3.2.1 *Translation.* A local test for translatory change checks directly whether the four colors found at an X-junction are consistent with a translatory change. Refer to figure 13. The two colors $\boldsymbol{a}$ and $\boldsymbol{b}$ must differ, as must the two colors $\boldsymbol{c}$ and $\boldsymbol{d}$. Furthermore, the difference in color $\boldsymbol{v}_1 = \boldsymbol{c} - \boldsymbol{a}$ must point in approximately the same direction as the difference in color $\boldsymbol{v}_2 = \boldsymbol{d} - \boldsymbol{b}$. We formulate this last constraint as follows:

$$\frac{\boldsymbol{v}_1 \cdot \boldsymbol{v}_2}{|\boldsymbol{v}_1||\boldsymbol{v}_2|} \geqslant s. \tag{5a}$$

The left-hand side is the dot product of the two difference vectors, normalized by the lengths of the two difference vectors. This is the cosine of the angle between the two vectors. The rule thus specifies that the cosine of the angle between the two vectors must exceed some threshold s or, equivalently, that the angle between the two vectors must be less than some corresponding threshold t:

$$\angle(\boldsymbol{v}_1, \boldsymbol{v}_2) \leqslant t. \tag{5b}$$

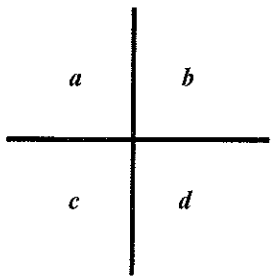


Figure 13. Four surfaces at an X-junction with tristimulus value vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, and $\boldsymbol{d}$. See text for discussion.

We performed Monte Carlo simulations of the changes in the lights from pairs of surfaces to validate this rule. Simulated reflectance functions of 1024 surfaces were generated randomly with the use of the nine-dimensional Fourier stochastic linear model (Appendix A). These reflectance functions were then multiplied, wavelength by wavelength, by CIE daylight illuminant D_{65} (Wyszecki and Stiles 1982). Pairs of spectral distributions from this set of simulated lit surfaces were then drawn from this set. For each of 10 000 pairs, chromatic change vectors were calculated for each member surface when seen in plain view and when seen under one of three transparent filters. We used single-pass filtering by broadband (Corning 1-57 glass), orange (Schott FG9 glass), and blue (Kodak Wratten 80C) transmission spectra (D'Zmura 1992) in these calculations. The vectors were either three-dimensional vectors of tristimulus values in the CIE XYZ space or two-dimensional vectors of chromaticities in the CIE (x, y) chromaticity plane. For each pair of surfaces, we recorded the angle between the three-dimensional tristimulus difference vectors and the angle between the two-dimensional chromaticity difference vectors.

We plot the results in figure 14. For each filter, the dark curve shows the cumulative frequency of angles between two-dimensional chromaticity differences, and the light curve shows the cumulative frequency for three-dimensional tristimulus value differences. Most of the angles are less than 30°, and almost all are less than 45°.

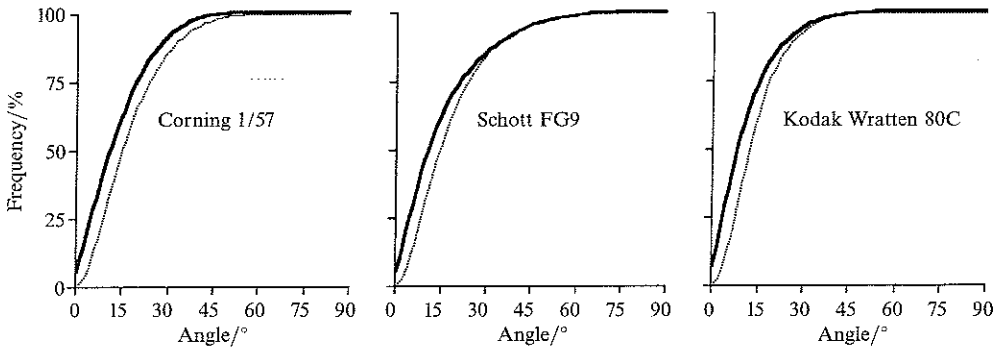


Figure 14. Cumulative frequencies of the angles between the chromatic change vectors for three filters (Corning 1-57 glass), orange (Schott FG9 glass), and blue (Kodak Wratten 80C) found in either the CIE XYZ color space or the CIE (x, y) chromaticity plane. From left to right: the cumulative frequencies for the chromaticity difference vectors (dark curves) found at 30° are 93.1%, 83.4%, and 89.4%; and at 45° are 98.7%, 94.4%, and 98.6%. The cumulative frequencies for the tristimulus value difference vectors (light curves) found at 30° are 92.3%, 83.5%, and 84.5%; and at 45° are 99.5%, 96.5%, and 97.3%.

The results suggest that the rule for determining whether colors at an X-junction are consistent with translation [equation (5b)] works best with a threshold in the range 30° – 45°.

The simplest global test for translatory change is to determine the average color change $\hat{t}$ across the contours of a candidate circuit and then check whether the change found at each piece of contour lies in the same direction as the average translation $\hat{t}$:

$$\frac{\boldsymbol{v}_i \cdot \hat{\boldsymbol{t}}}{|\boldsymbol{v}_i||\hat{\boldsymbol{t}}|} \geqslant s, \quad \forall i. \tag{6}$$

This rule can be used to detect a circuit with a uniform color translation and provides an estimate of this translation.

The global test fails to detect heterogeneous translations, however. Figure 15 (color plate 1, page 475) shows an example. The circular overlay in this figure is an equiluminous translation that varies in hue as a function of polar azimuthal angle, measured relative to the center of the figure, to create a color circle. The translation varies in saturation as a function of distance from the center; it ramps linearly between zero at the center and its maximal value at the circle border.

The average translation around the circuit of the transparent region in figure 15 is zero, and the global rule [equation (6)] fails miserably to detect the transparent region. Yet the local rule [equation (5a) or (5b)] works well to detect the region, because the color change at each X-junction is consistent with translatory change. This suggests that one should use the local rule to detect transparent regions and a global procedure to identify the parameters of a translation for use in color separation.

3.3.2.2 *Convergence.* A local test for convergent motion at an X-junction is to check whether the distance $|\boldsymbol{a} - \boldsymbol{b}|$ between two colors in plain view at an X-junction (see figure 13, facing page) is greater than the distance $|\boldsymbol{c} - \boldsymbol{d}|$ between the corresponding colors on the potentially transparent side of the X-junction.

There are two ways to estimate globally the parameters of uniform convergence. In the first the amount of convergence and the target of convergence are estimated sequentially, while in the second they are estimated simultaneously.

This figure is on color plate 1, page 475

Figure 15. Local but not global coherence: changing color filter properties. See text for discussion.

The first way depends on noting that the sum of the original color vector $\boldsymbol{a}_1$ of a surface and a certain amount $\hat{\gamma}$ of the difference $\boldsymbol{v}_1 = \boldsymbol{b}_1 - \boldsymbol{a}_1$ between the new and the original color vectors of this surface equals the point $\boldsymbol{g}$ of convergence:

$$\boldsymbol{a}_1 + \hat{\gamma}\boldsymbol{v}_1 = \boldsymbol{g}. \tag{7}$$

Because this holds true for all surfaces along the border of the transparent area, we can require that the squared difference between each surface's expression for $\boldsymbol{g}$, for each possible pair of surfaces, be minimal:

$$\frac{\partial}{\partial\hat{\gamma}} \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{3}(a_{ik} + \hat{\gamma}v_{ik} - a_{jk} - \hat{\gamma}v_{jk})^2 = 0, \tag{8}$$

where i and j run over the number n of surfaces and k runs over the three dimensions of each color vector. This minimization provides a global estimate $\hat{\gamma}$ of the scalar in equation (7), which can then be used to determine globally an estimate $\hat{\boldsymbol{g}}$ of the point of convergence.

A second method uses a singular value decomposition (Press et al 1988) to estimate simultaneously the parameters of a uniform convergence. Suppose that there are n surfaces along the border of the candidate transparent area that are seen both in plain view and through the filter. Then for each surface we have from equation (3):

$$\boldsymbol{b}_i = (1-\alpha)\boldsymbol{a}_i + \alpha\boldsymbol{g}, \quad i = 1, ..., n, \tag{9a}$$

which may be rewritten as follows:

$$\boldsymbol{b}_i = \beta\boldsymbol{a}_i + \boldsymbol{g}', \quad i - 1, ..., n. \tag{9b}$$

The latter is a set of simultaneous linear inhomogeneous equations for which a singular value decomposition can provide a least-squares solution. The equations generalize the model of Gerbino et al (1990) from the case of two achromatic surfaces to the case of arbitrarily many color surfaces. One may rewrite equation (9) in matrix-vector form to make clear the applicability of the singular value decomposition (Press et al 1988):

$$\begin{bmatrix} a_{11} & 1 & 0 & 0 \\ a_{12} & 0 & 1 & 0 \\ a_{13} & 0 & 0 & 1 \\ & \vdots & & \end{bmatrix} \begin{bmatrix} \beta \\ g_1' \\ g_2' \\ g_3' \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \\ \vdots \end{bmatrix}. \tag{10}$$

in which the numbers a_{ij} ($i = 1, ..., n; j = 1, 2, 3$), are the tristimulus values of the ith surface when seen through the transparent filter. Two such surfaces suffice to provide more equations ($3n$) than unknowns (4).

This method can be used to recover the parameters of an arbitrary uniform convergence (or, as a special case, an arbitrary uniform translation; see section 2.2.2). It is applied to the chromatic changes along the border of a candidate transparent region to determine the extent to which the changes are coherent and to estimate the parameters of the changes. The major drawback of this method is its reliance on the assumption of filter uniformity—it cannot handle cases like that in figure 15.

3.4 Separation of transparent overlay from underlying surfaces

Figure 16 (color plate 3, page 482) shows the separation of the input image into representations of the chromatic properties of the transparent overlay (figure 16a) and the surfaces (figure 16b). To recover the colors of surfaces that are seen both in plain view and through the overlay, the algorithm simply propagates the plain-view chromatic properties into the region of transparency. For surfaces that lie entirely within the region of transparency, the algorithm uses the estimate $\hat{\boldsymbol{t}}$ of the translation, found

for this sample image, to undo the change in surface color. Recall that the effect of the translation is to add a translation $\boldsymbol{t}$ to the original color $\boldsymbol{a}$ to create the observed color $\boldsymbol{b}$ seen through the overlay [equation (1)]. The estimate $\hat{\boldsymbol{a}}$ of the original surface chromatic properties is thus found through the inverse operation: subtracting the estimated translation from the observed color:

$$\hat{\boldsymbol{a}} = \boldsymbol{b} - \hat{\boldsymbol{t}}. \tag{11}$$

The separation in the case of convergence proceeds in the same manner. The chromatic properties of surfaces found in plain view are propagated within the transparent region. For surfaces that lie entirely within the transparent region, estimates of the convergence parameters are used to invert the convergence. The global estimates $\hat{\beta}$ and $\hat{\boldsymbol{g}}'$ of the 'amount' and 'target' of convergence, respectively [see equations (9) and (10)], are used together with the observed color $\boldsymbol{b}$ of equation (2) to provide an estimate $\hat{\boldsymbol{a}}$ of the original color:

$$\hat{\boldsymbol{a}} = \frac{\boldsymbol{b} - \hat{\boldsymbol{g}}'}{\hat{\beta}}. \tag{12}$$

This rule works well to reconstitute surface color in regions of uniform chromatic convergence. Note that in cases of translation the estimated parameter $\hat{\beta}$ is close to one, and the procedure for convergence [equation (12)] reduces to that for translation [equation (11)].

This figure is on color plate 3, page 482

Figure 16. Algorithm: Separation of transparent overlay from surfaces for the image in figure 11a. (a) The filter. (b) The surfaces. See text for discussion.

4 Discussion

Perceiving a transparent region depends on the visual detection and identification of a process that causes lights from surfaces to change systematically. We have found that the systematic changes in color to which observers are sensitive are translation and convergence.

Translation corresponds to what had earlier been identified as subtractive color mixture by Beck (1978) and multiplicative transparency by Kersten (1991). Each of these appellations is associated with a physical model of transparency, even though physical transparency is neither a necessary nor a sufficient condition for transparency perception (Metelli 1974). Translation in color space, on the other hand, is expressed directly in terms of human color perception. Translations provide a first-order approximation to a wide range of multiplicative filters and generalize this class of filtering to include physically unrealizable cases like equiluminous color shifts.

Convergence corresponds to what had earlier been identified as additive color mixture by Metelli (1974) and others. The additive mixtures of two colors span a plane in a standard three-dimensional color space. To generate colors on the line segment joining two colors, one must simultaneously add light to one color and subtract light from the other [equation (2)], ie for transparency one requires highly constrained additive color mixtures. Viewing such mixtures in a three-dimensional color space suggests immediately the more informative term convergence.

We readily perceive compositions of translation and convergence, which are again convergences, as transparent (figures 6–8, color plate 2, page 479). The parameters that describe such convergences are two: the colour vector of the center of convergence, and the scalar degree α of convergence [equation (4)]. It is of some interest to determine whether

human observers can distinguish between these parameters when making judgments of color transparency. In many psychophysical studies of transparency, only one form (translation) or another (convergence) of transparency is considered, and joint judgments of transparency parameters are not elicited. However, Gerbino and colleagues (1990) have shown that observers make settings in an achromatic matching task which are consistent with a scalar version of equation (9b) that has a multiplicative constant and an additive constant. Study of the color figures (eg figures 6–8, color plate 2, page 479) suggests that one can estimate both types of chromatic parameters; one awaits an experimental characterization of human performance in such an estimation task.

The perception of transparency with equiluminous shifts shows that systematic changes in luminances are not required for transparency to be perceived. It is important to note that there is no common, natural process that produces equiluminous translations. For a light to remain equiluminous when shifted, one must both subtract light in one region of the spectrum and add light in another region of the spectrum to balance the luminance. If one has several surfaces with hues that span the spectrum, then it is impossible for either episcotisters or natural filters with spatially uniform properties to produce such equiluminous shifts. There is no ready explanation in terms of multiplicative models or subtractive color mixture. Clearly, one must base the study of transparency perception on perception itself, rather than on physical models.

By representing systematic chromatic change in terms of vector fields in color space we have been led to examine stimuli with a larger than typical number of surfaces. Most studies of transparency perception work with a total of four colors, created by a putative filter that overlies two surfaces (eg Metelli 1974; Beck et al 1984; Fukuda and Masin 1994). This restricted situation can give rise to questions that are more readily answered when one considers more complex stimuli. For instance, Metelli's estimates of transparency are based on an assumption of 'balanced' transparency, wherein the degree of transparency [convergence parameter α, equation (2)] is assumed to be constant for all surfaces. The visual system makes no such assumption of homogeneity (Fukuda and Masin 1994): using many surfaces one can easily demonstrate that the transparent overlay is free to vary (figure 15, color plate 1, page 475). The real problem is finding how the visual system separates a heterogeneous overlay from underlying surfaces.

The detection and separation algorithm presented here is designed primarily to illustrate the role of translation and convergence in color space. Its spatial processing is not sophisticated. In particular, the region growing scheme that is used for contour and intersection determination is weak. A first problem is the inability to segment properly images that include noise or, more generally, chromatic change across single surfaces. Second is the divergence between this sort of scheme and current knowledge concerning human visual processing, which analyzes color images at multiple spatial scales (Watson 1987; D'Zmura 1997). We hope to address this problem in future work by using a more robust scheme for color image segmentation.

The algorithm does have several strengths. It can detect transparent regions that lie along the image border through consideration of T-junctions. This is accomplished through the inclusion of T-junctions along the image border within the XJAG. Indeed, there are two arguments in favor of retaining all T-junctions in the determination of the contour of a transparent region. The first is the possible presence of implicit X-junctions (Watanabe and Cavanagh 1993). A second argument is the apparent ability of T-junctions to terminate the contour of a transparent region within the center of an image. Consider the curl example in figure 9a (color plate 1, page 475). The blue-green change at the top left of the central square region is an accidental translation which forms a small region of transparency. This region appears to be bounded below by at least one T-junction, namely that between the blue, green, and

gray surfaces in the center of the image. We eliminated 'central' T-junctions in our algorithm to reduce the number of possible circuits that must be considered; central T-junctions are readily admitted back into the coherence computation, however.

Another strength of the algorithm is its ability to detect heterogeneous transparent regions like that in figure 15 (color plate 1, page 475). The local tests for translation and convergence make this possible. Furthermore, the procedures for estimating the global parameters of translation and convergence let us perform a color separation of transparent overlay and underlying surfaces when the overlay is spatially uniform. The most interesting problem that remains to be solved, again, is how to separate spatially nonuniform overlays from underlying surfaces.

Acknowledgements. We thank Alan Gilchrist, Anya Hurlbert, Donald MacLeod, John Mollon, Lothar Spillmann, Paul Whittle, and an anonymous referee for their helpful comments. A preliminary version of this work was presented at the 'Colour in Context' conference of the Colour Society in Durham, England, April 1996. This work was supported by National Eye Institute grant EY10014 and by a Bourse d'Accueil from the Région Rhône-Alpes to M D'Zmura.

References

Beck J, 1978 "Additive and subtractive color mixture in color transparency" *Perception & Psychophysics* **23** 265–267

Beck J, Prazdny K, Ivry R, 1984 "The perception of transparency with achromatic colors" *Perception & Psychophysics* **35** 407–422

Brainard D, 1995 "Colorimetry", in *Handbook of Optics* Volume 1 *Fundamentals, Techniques, and Design* Ed. M Bass (New York: McGraw-Hill) pp 26.1–26.54

Brill M H, 1984 "Physical and informational constraints on the perception of transparency and translucency" *Computer Vision, Graphics and Image Processing* **28** 356–362

Craven B J, Foster D H, 1992 "An operational approach to colour constancy" *Vision Research* **32** 1359–1366

Da Pos O, 1989 *Trasparenze* (Padua: Icone)

Derrington A M, Krauskopf J, Lennie P, 1984 "Chromatic mechanisms in lateral geniculate nucleus of macaque" *Journal of Physiology (London)* **357** 241–265

D'Zmura M, 1992 "Color constancy: surface color from changing illumination" *Journal of the Optical Society of America A* **9** 490–493

D'Zmura M, 1997 "Color contrast gain control", in *Colour Vision* Eds W Backhaus, R Kliegl, J Werner (Berlin: Walter de Gruyter) in press

D'Zmura M, Iverson G, 1997 "A formal approach to color constancy: the recovery of surface and light source spectral properties using bilinear models", in *Recent Progress in Mathematical Psychology* Eds C Dowling, F Roberts, P Theuns (Mahwah, NJ: Lawrence Erlbaum Associates) in press

D'Zmura M, Iverson G, Singer B, 1995 "Probabilistic color constancy", in *Geometric Representations of Perceptual Phenomena* Eds R D Luce, M D'Zmura, D D Hoffman, G Iverson, K Romney (Mahwah, NJ: Lawrence Erlbaum Associates) pp 187–202

D'Zmura M, Lennie P, 1986 "Mechanisms of color constancy" *Journal of the Optical Society of America A* **3** 1662–1672

D'Zmura M, Mangalick A, 1994 "Detection of contrary chromatic change" *Journal of the Optical Society of America A* **11** 543–546

D'Zmura M, Singer B, 1996 "The spatial pooling of contrast in contrast gain control" *Journal of the Optical Society of America A* **13** 2135–2140

Foster D H, Craven B J, Sale E R H, 1992 "Immediate colour constancy" *Ophthalmological and Physiological Optics* **12** 157–160

Fukuda M, Masin S C, 1994 "Test of balanced transparency" *Perception* **23** 37–43

Gerbino W, Stultiens C I F H J, Troost J M, Weert C M M de, 1990 "Transparent layer constancy" *Journal of Experimental Psychology: Human Perception and Performance* **16** 3–20

Gibson J J, 1979 *The Ecological Approach to Visual Perception* (Boston, MA: Houghton Mifflin)

Gonzalez R C, Wintz P, 1977 *Digital Image Processing* (Reading, MA: Addison-Wesley)

Hogg R, Craig A, 1978 *Introduction to Mathematical Statistics* 4th edition (New York: Macmillan)

Kersten D, 1991 "Transparency and the cooperative computation of scene attributes", in *Computational Models of Visual Processing* Eds M S Landy, A J Movshon (Cambridge, MA: MIT Press) pp 209–228

Koenderink J J, 1986 "Optic flow" *Vision Research* **26** 161–180

Krauskopf J, Williams D R, Heeley D M, 1982 "The cardinal directions of color space" *Vision Research* **22** 1123–1131
MacLeod D I A, Boynton R M, 1979 "Chromaticity diagram showing cone excitation by stimuli of equal luminance" *Journal of the Optical Society of America* **69** 1183–1186
Maloney L T, Wandell B A, 1986 "Color constancy: a method for recovering surface spectral reflectance" *Journal of the Optical Society of America A* **3** 29–33
Metelli F, 1974 "The perception of transparency" *Scientific American* **230**(4) 91–98
Mulligan J B, 1993 "Nonlinear combination rules and the perception of visual motion transparency" *Vision Research* **33** 2021–2030
Pavlidis T, 1980 *Structural Pattern Recognition* (Berlin: Springer)
Press W H, Flannery B P, Teukolsky S A, Vetterling W T, 1988 *Numerical Recipes in C* (New York: Cambridge University Press)
Russ J C, 1995 *The Image Processing Handbook* 2nd edition (London: CRC Press)
Schettini R, 1993 "A segmentation algorithm for color images" *Pattern Recognition Letters* **14** 499–506
Schey H M, 1992 *Div, Grad, Curl and All That* 2nd edition (New York: Norton)
Shafer S, 1985 "Using color to separate reflection components" *Color Research and Application* **10** 210–218
Singer B, D'Zmura M, 1994 "Color contrast induction" *Vision Research* **34** 3111–3126
Singer B, D'Zmura M, 1995 "Contrast gain control. A bilinear model for chromatic selectivity" *Journal of the Optical Society of America A* **12** 667–685
Smith V C, Pokorny J, 1975 "Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm" *Vision Research* **15** 161–171
Tudor-Hart B, 1928 "Studies in transparency, form and color" *Psychologische Forschung* **10** 255–298
Watanabe T, Cavanagh P, 1993 "Transparent surfaces defined by implicit X junctions" *Vision Research* **33** 2339–2346
Watson A B, 1987 "The cortex transform: rapid computation of simulated neural images" *Computer Vision, Graphics, and Image Processing* **39** 311–327
Weert C M M de, 1986 "Superimposition of color information" *Color Research and Application* **11** 21–26
Wyszecki G, Stiles W S, 1982 *Color Science. Concepts and Methods, Quantitative Data and Formulae* (New York: Wiley)

APPENDIX A

We apply a stochastic linear model (D'Zmura et al 1995) to generate surface reflectance functions for use in simulated images (eg figure 2a, color plate 1, page 475). Such a reflectance function is generated by a model that combines linearly the first nine terms in a Fourier series analysis of functions on the interval [400, 700] nm of visible wavelengths. The model has the form:

$$\boldsymbol{R}(x) = \boldsymbol{c}_0 + \sum_{f=1}^{4} (\boldsymbol{a}_f \cos 2\pi f x + \boldsymbol{b}_f \sin 2\pi f x), \tag{A1}$$

where x is a wavelength variable that has been transformed to range linearly from 0 to 1 over the interval 400 to 700 nm, respectively: $x = (\lambda - 400)/(700 - 400)$, for $\lambda \in [400, 700]$ nm.

The nine random coefficients $\boldsymbol{c}_0$, $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$, $\boldsymbol{a}_4$, $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{b}_3$, and $\boldsymbol{b}_4$ in the expansion (A1) are chosen in a way to generate only physically realizable reflectance functions, namely functions that take on values only in the interval [0, 1]. They are chosen in order of increasing frequency. The average term $\boldsymbol{c}_0$ is beta distributed. The beta distributions form a family with two parameters that control the mean and variance of the distribution (Hogg and Craig 1978). A beta distribution is nonzero only on the interval [0, 1], so that the (constant) reflectance functions generated by this term of the model alone are physically realizable.

Having sampled the beta distribution to generate a particular value c_0, one then selects the amplitude $c_1 = (a_1^2 + b_1^2)^{1/2}$ of the sinusoidal component at frequency 1 from a distribution that takes on values only on the interval $[0, \min(c_0, 1 - c_0)]$. One such distribution is the uniform distribution on the interval $[0, \min(c_0, 1 - c_0)]$. Yet what is important is the interval, which ensures that the entire range of saturations can be generated, given the constraint of physical realizability and the prior choice of c_0. The phase of the component at frequency 1 is chosen from the uniform distribution on $[0, 2\pi]$.

Coefficients for frequencies 2, 3, and 4 are found in the same way. Having chosen amplitudes c_0 through c_{i-1}, one chooses the amplitude c_i from a distribution that takes on values only in the (remaining) interval $\left[0, \min \sum_{j=0}^{i-1} c_j, 1 - \sum_{j=0}^{i-1} c_j\right]$; again, the phase is chosen randomly.