Surface decomposition accompanying the perception of transparency

TAKEO WATANABE*†1,2 and PATRICK CAVANAGH1

1Department of Psychology, Harvard University, Cambridge, MA, USA
2Visual Science Laboratory, Arizona State University West, Phoenix, AZ, USA

Received for publication 1 June 1992

Abstract—At the retina, each location can have only one value of luminance or color. When transparency is perceived, however, different surface qualities can be redistributed to two or more apparently superimposed layers. The experiments described here explored the characteristics of this surface decomposition. It is shown that the surface decomposition occurs rapidly, it affects even early stages of visual processing, and it involves attributes such as texture and motion as well as color and brightness. In the first experiment, the recognition advantage for transparent overlapping digits demonstrated that the surface decomposition accompanying transparency occurs within 60 ms. In the second, the separation of overlying, orthogonal grids due to surface decomposition was found to influence the strength of the McCollough effect, an effect attributed to early cortical processing. Finally, when a transparent surface appears to extend over areas that are physically identical to the background, qualities of the transparent overlay such as texture and motion, as well as color or brightness (e.g., the neon color effect) appear to spread to the illusory overlay.

1. INTRODUCTION

Most computational models of vision assume that external world is opaque and therefore that each location of the visual field is represented by only one value for each surface attribute such as brightness and color. However, in Fig. 1a, we see two overlapping squares with one dark square visible through another lighter square. This phenomenon is called subjective transparency (e.g., Metelli, 1974; Beck et al., 1984). Within the overlapping region shared by the two squares, both surface qualities, light and dark, may be perceived simultaneously and appear to be assigned to the separate depth levels of the two surfaces. This indicates that, contrary to the basis of simple computational models, the visual system has the ability to decompose an image value at a single point into several, superimposed perceived values.

In recent work, we have explored the characteristics of this decomposition. In the experiments which will be reviewed here we report that the surface decomposition is fast, it affects even early stages of visual processing, and it involves attributes such as texture and motion as well as color and brightness.

2. SURFACE DECOMPOSITION IS FAST

We were first interested in measuring the speed with which the superimposed surfaces become separately available. At this point, the image of Fig. 1 is organized as two separate squares rather than three adjacent regions. One direct method to determine the moment at which the squares 'emerge' as separate shapes is to test the recognition

*To whom all correspondence should be addressed.
†Present address: Vision Science Laboratory, Arts and Sciences, Arizona State University West, 4701 W. Thunderbird Road, PO Box 37100, Phoenix, AZ 85069-7100, USA.
of these shapes. Rather than shapes, however, we used overlapping digits in a method developed by de Weert (1986).

Before describing this stimulus, we will briefly review the image conditions which lead to the perception of transparency. Metelli (1974) and Beck et al. (1984) have shown that observers report an impression of overlapping transparent surfaces in a flat image like Fig. 1a as long as the light values in the image correspond to the following two simple constraints derived from the physics of transparency: (1) direction of the lightness contrast across an opaque border does not change in the transparent region; and (2) the lightness difference across an opaque border must be reduced in the transparent region. If one of these constraints is violated, it is difficult to see transparency. In Fig. 1b, for example, the first rule is violated.

In Fig. 2a, two digits overlap two others and the luminance combinations of the four digits and the overlapping areas are appropriate to simulate physical transparency. The reader may see the lower two transparent digits as overlying the upper two opaque digits. When organized as overlying transparent surfaces, the overlapped and nonoverlapped regions of each digit appear to group together facilitating their recognition. In Fig. 2b, the luminance combination is not valid for transparency. The perceptual separation of the individual digits is much less compelling in this case and they may be harder to recognize in a brief presentation.

In order to evaluate the minimum duration for which the separation of the digits was available, we used a pattern identification task. De Weert (1986) presented two overlapping colored five-letter words and asked observers to name both of them. Each word was presented in a different color and the overlapping regions were presented in a third color. De Weert found that the lowest error rates in naming the words, and therefore presumably the highest degree of transparency occurred when the overlap color (yellow) was the additive mixture of the two word colors (red and green).

In our first experiment (Watanabe and Cavanagh, 1991b, 1992), we used de Weert's pattern identification task but varied luminance rather than color to evaluate the shortest duration that produced a recognition advantage for stimuli with luminances
Figure 2. Examples of stimuli used in: (a) the valid luminance condition, (b) the invalid luminance condition, and (c) the silhouette condition of Experiment 1. The luminances of the stimuli shown here are representative of those used in the stimuli but do not reproduce them exactly.
valid for transparency. To start with, we examined how well the pattern identification results correlated with subjective transparency.

**Experiment 1**

Experiment 1 verified that the facilitation of identification was linked to subjective reports of transparency. As a test stimulus, four randomly selected digits were presented simultaneously (120 ms), as shown in Fig. 2. They were followed by a mask consisting of white and black rectangles. The luminance of the digits varied in three conditions: (a) a valid luminance condition in which the luminance combination was valid for subjective transparency; (b) an invalid luminance condition where it was not; and (c) a silhouette condition in which all digit regions had the same luminance. In the valid luminance condition, the luminances of the overlapping areas, the remaining parts of the upper digits, and the lower digits were 30, 22 and 26 cd m$^{-2}$, respectively. This luminance combination is valid for physical transparency. In the invalid luminance condition, these same areas had luminances of 22, 30, and 26 cd m$^{-2}$, respectively. This combination violates Metelli’s first constraint according to which a transparent surface cannot change the order of the lightness values across a contour it covers. In the silhouette condition, these areas were all 26 cd m$^{-2}$. The background in each of the three conditions was 17 cd m$^{-2}$ in luminance. The average Michelson luminance contrasts between overlapping area and the remaining parts of the upper and lower digits were 11%, 12% and 0% in the valid, invalid, and silhouette conditions, respectively. The average luminance contrasts between digits and the background were 17%, 21%, and 24% in the three conditions, respectively. A whole session consisted of two blocks: the pattern identification task and the rating task for the degree of transparency. In the pattern identification task, subjects identified as many as possible of the briefly presented 4 test digits. In the rating task, the subjects rated the degree of transparency of the same stimuli as in the pattern identification task with 0 indicating no transparency and 10, a very strong impression of transparency.

Figure 3 shows the mean value in the rating task and the mean percentage of correct responses in the pattern identification task (both for the same five subjects), plotted for each of the three luminance conditions.

These results showed that high performance in the pattern identification task was accompanied by high ratings of transparency. The identification advantage for the transparent stimuli is likely a result of surface separation that phenomenally accompanies transparency. That is, if transparency occurs, the regions of the stimuli are appropriately grouped as separate digits and the digits are more easily identified. If transparency does not occur, the stimuli appear to be somewhat a jumble of bits and pieces and the digits are more difficult to identify.

**Experiment 2**

In Experiment 2, again using the pattern identification task, we examined the minimum stimulus duration required for the perception of transparency (Watanabe and Cavanagh, 1991b, 1992).

As shown in Fig. 4, there were four stimulus conditions: (a) valid luminance and digits-overlapped, (b) invalid luminance and digits-overlapped, (c) ‘low’ luminance and digits-separated, and (d) ‘high’ luminance and digits-separated. The luminance combinations in the luminance valid and invalid conditions were identical to those in
Experiment 1. The layout of digits in the digits-overlapped conditions was identical to that in Experiment 1. The digits in the digits-separated conditions were presented with no spatial overlap. In the ‘valid’ luminance condition, the luminance in the upper two digits was 22 cd m$^{-2}$ and that in the lower two digits was 26 cd m$^{-2}$, the same values as for the valid overlapped condition except that there is no overlapped region. In the ‘invalid’ luminance condition, the luminance in the upper two digits was 30 cd m$^{-2}$ and that in the lower two digits was 26 cd m$^{-2}$, the same as those in the invalid overlapped condition. The test stimulus duration was varied randomly in each trial and the test was followed immediately by a mask.

In Fig. 5, the mean percentages of correct responses in the four conditions are plotted as a function of duration. The percentages of correct responses in the separated conditions did not vary as a function of the luminance (‘valid’ or ‘invalid’) and so are combined in the graph. The data show that with as little as 60 ms duration, the percentage of correct responses is higher in the valid luminance condition than in the invalid luminance condition. Secondly, even at 150 ms duration, the percentage of correct responses in the valid luminance condition is still lower than in the condition with separated digits.

These results suggest that the performance advantage seen to accompany the perception of transparency in Experiment 1 occurs for a stimulus presentation as short as 60 ms and, therefore, that the surface separation process is very fast and automatic. On the other hand, even at large intervals the perceptual separation of the transparent digits is not as effective as physical separation.

The results obtained in Experiments 1 and 2, however, may have an alternative,
low-level explanation that does not involve transparency explicitly. As is evident in
the nontransparent cases in Figs 1b, 2b, and 4b, the direction of contrast across image
contours reverses wherever one contour intersects the other. These contrast reversals
may interfere with the encoding of the image contours for these images. Low-level
physiological mechanisms which are orientation-selective respond best when a con-
tour is light on one side and dark on the other (Hubel and Wiesel, 1962). They would
respond poorly to a contour which reversed contrast in the middle of the receptive
field. The image in the valid (transparent) conditions of Experiment 1 and 2 had no
contrast reversals at contour intersections and so would suffer less interference than
the invalid conditions where contrast reversed at each intersection. In a preliminary
experiment ($n = 1$), we added a new condition, in which the luminance of the
overlapping area of the digits is too high for transparency (violating Metelli’s second
law), to the four conditions in Experiment 2. In this case, there was no contrast
reversal. We found the performance in this condition was significantly lower than that
in the valid luminance and overlapping condition but was not significantly different from that in the invalid luminance and overlapping condition. The results suggest that the higher performance in Experiment 2 is due to perception of transparency rather than to the same direction of contrast across image contours.

3. SURFACE DECOMPOSITION AFFECTS EARLY VISION

In the pattern identification task, we obtained results suggesting that the surface decomposition involved in transparency perception occurs quite rapidly. In this section, we examine whether the separation affects early stages of visual processing. Specifically, we used the orientation-contingent color aftereffect originally studied by McCollough (1965), an effect which has been suggested to occur at an early stage in the visual information processing.

In this paradigm, an observer views a green and black, vertically striped pattern alternating every few seconds with a red and black horizontally striped pattern. Following several minutes of adaptation to these stimuli, a test stimulus made up of white and black vertical stripes appears to be tinged in red whereas a test stimulus of white and black horizontal stripes appears to be tinged in green. That is, the test stripes appear in a color complementary to the color of the adaptation of the same orientation. The McCollough effect is based on orientation, suggesting that its site must at least be cortical where the first orientation-specific units arise (Stromeyer, 1978; Dodwell and O'Shea, 1987; Livingstone and Hubel, 1987). However, the effect is also largely monocular (McCollough, 1965; Murch, 1972; Stromeyer et al., 1973; MacKay and MacKay, 1975; Skowbo et al., 1975; White and Riggs, 1975), suggesting a site that precedes the emergence of binocular units which also occurs in the first area of the visual cortex (Hubel and Wiesel, 1962). These two results are somewhat conflicting but imply that the locus of the McCollough effect is probably in the earliest
stages of cortical visual processing (also see, Humphrey et al., 1991; Savoy and Gabrieli, 1991).

An interaction of the surface decomposition accompanying transparency with the McCollough effect will therefore indicate that transparency can influence visual processing at a very early stage. This research is described in more detail in Watanabe et al. (1992).

**Experiment 3**

Four test patterns were used, as shown in Fig. 6. One is the valid transparency pattern (Fig. 6a) consisting of small squares of four different luminances set so that the vertical and horizontal stripes appear to overlie each other transparently. When the luminances were set to be invalid for transparency, the horizontal and vertical structure appeared to be replaced by a checkerboard-like organization of individual squares. The luminance of the squares constituting the isolated sections of the vertical and horizontal bars was 50.0 cd m⁻². The luminances of the intersections and the background were 75.0 and 0.5 cd m⁻², respectively. The invalid transparency patterns are the same as the valid transparency pattern except that the luminance of the intersection square is too low (25 cd m⁻²) as in Fig. 6b or too high (125 cd m⁻²) as in Fig. 6c to simulate transparency. In the remaining test pattern (Fig. 6d), two vertical and two horizontal gratings are presented in physically separate areas with no overlap. The luminances of bars and the background were the same as those in Fig. 6a-c. Before the experiment, it was confirmed that the valid pattern appeared as a transparent overlay of grids of horizontal and vertical stripes for all the subjects while the invalid patterns did not. The result of pretest showed that none of the subjects (n = 6) had any residual orientation contingent color aftereffect prior to the experiment.

In the experiment, initially the subjects were exposed to the two adapting stimuli, vertical red stripes and horizontal green stripes, alternating at 4-s intervals. The total inspection time was 10 min. To prevent the systematic buildup of negative afterimages, the subjects were instructed not to fixate any particular region of the adapting stimuli, but rather to let their gaze wander around the colored gratings. Immediately after adaptation, testing began.

At first, the test pattern whose intersections were too dark for transparency (Fig. 6c) was presented. The initial red/green saturation of the vertical and horizontal elements was determined randomly. The subjects’ task was to adjust the saturation of the test pattern until it appeared uniformly achromatic (color cancellation technique e.g., Riggs et al., 1974; Webster et al., 1987). Immediately after the subject clicked the mouse to indicate an acceptable setting, the adaptation stimuli were presented again for 16 s (the two adaptation stimuli alternating in 4-s intervals). A new test pattern was then presented (the valid transparency pattern, Fig. 6a) and the setting procedure repeated. The same top-up of adaptation, followed by test, continued for the remaining two test patterns (grid with intersections too bright for transparency, Fig. 6c, and the test with physically separated horizontal and vertical gratings, Fig. 6d) and the four-test cycle continued for a total of 20 repetitions. Thus, the whole experimental session consisted of 80 (four test patterns and 20 repetitions) settings and lasted about 90 min.

The mean of excitation purities of red and green added to the test stimulus to cancel the apparent colors induced by adaptation in the experiment was taken as a measure
Figure 6. Depiction of the test figures. The intersections were: (a) valid; (b) too dark; and (c) too bright for transparency. In (d), vertical and horizontal stripes were physically separated. These figures demonstrate the differences among the test figures but are not exact reproductions of them (see text for detailed descriptions).

of the strength of adaptation. The excitation purity of each segment was defined as the ratio of the distance between the display white point (0.33 and 0.33 in the CIE x- and y-chromaticity coordinates) and the CIE chromaticity coordinates of the color to the distance of the dominant wavelength of the color from the white point. Figure 7 shows the mean excitation purity necessary to cancel the apparent colors for the invalid (too dark) test pattern, valid test pattern, invalid (too bright) test pattern, and physically separated pattern. The mean excitation purity was significantly higher for
the valid luminance combinations than those for invalid (too dark and too high) luminance combinations. This was consistent throughout all of the six subjects.

This result suggests that the McCollough effect is influenced by surface decomposition accompanying the perception of transparency. That is, the subjective organization of the grids in Fig. 6a as overlapping horizontal and vertical stripes was more similar to the physically separated adapting stripes than to the checkerboard pattern seen in Fig. 6b and c. On the other hand, the scores for the valid luminance combination were lower than those for the physically separated patterns for all subjects. This finding indicates, as in the results of Experiment 2, the surface decomposition accompanied by the transparency is not as effective as physical separation.

4. TRANSPARENCY INVOLVES TEXTURE AND MOTION

In this section, we will argue that the visual system can make up a subjective transparent surface where there is no stimulus information and that the subjective surface can have attributes not only of brightness and color but also texture and motion (see Watanabe and Cavanagh, 1991a).

The lines in the pattern shown in Fig. 8a (Ehrenstein, 1941) induce disk-shaped illusory surfaces within the area of missing intersections. When colored or gray crosses, which by themselves induce no interesting visual effect (Fig. 8b), are inserted
Figure 8. (a) The Ehrenstein figure—radially arranged lines induce illusory disks between their inner tips. (b) Gray crosses which, on their own, appear as gray crosses. (c) If the same gray crosses as in (b) are inserted into the gaps of the Ehrenstein figure, the brightness (and color, if different) of the crosses is seen to spread out to fill an illusory disk in a visual illusion known as the neon color spreading effect.
into the intersections of the Ehrenstein figure (Ehrenstein, 1941), we see transparently colored (or gray in Fig. 8b) disks through which completed white lines of underlying grids are seen. This visual illusion is known as the neon color spreading effect (van Tuijl, 1975; Redies and Spillmann, 1981; Watanabe and Sato, 1989; Watanabe and Takeichi, 1990). Many authors have suggested that there is a close relationship between the neon color spreading effect and transparency (Redies and Spillmann, 1981; Meyer and Senecal, 1983; de Weert and Kruysbergen, 1987; Grossberg, 1987; Nakayama and Shimojo, 1990; Ramachandran, 1990; Takeichi et al., 1992). The central feature of this phenomenon is that a surface attribute which is only visible when silhouetted against the grid line is assigned to the full extent of the subjective disk which appears to cover the intersection.

However, surfaces have attributes not only of color and brightness but also of texture, motion and depth. A percept of a transparent surface may therefore make these attributes spread over a subjective surface as well. Specifically, we (Watanabe and Cavanagh, 1991b) examined whether surface attributes of texture and/or motion can be assigned to areas where no texture or motion signal was present by replacing the gray crosses in Fig. 8c with ones consisting of static (in Experiment 4) or moving texture (in Experiment 5).

**Experiment 4**

Four figures were used: (1) eight crosses filled with a texture made up of black and white diagonal stripes were inserted in the gaps of the Ehrenstein figure (Fig 9a); (2) just the eight textured crosses were presented in isolation (Fig 9b); (3) and (4) were identical to (1) and (2) except that the diagonal texture was replaced with a texture of black and white dots (Fig. 9c and d). Subjects were asked to fixate the black cross in the center of the figures and to report whether the texture of the crosses appeared to spread outside of the cross or not. Table 1 shows the percentage of the subjects who saw the texture spread outside of the crosses as well as the percentage of those who saw a well-defined contour around the texture for each stimulus. All subjects who saw the contour reported that it was circular. For the striped and dotted texture crosses, the percentages of the subjects who saw the texture spreading were significantly higher when the textured crosses were embedded in the Ehrenstein figures than when they were presented alone ($z = 4.12, p < 0.01$ for striped texture and $z = 3.79, p < 0.01$ for dotted texture — comparisons of binomial probabilities).

**Experiment 5**

In Experiment 4, we found that texture appeared to spread outside of the crosses when placed in the Ehrenstein figure. In this experiment, the texture inside the crosses was made to move at 4 deg s$^{-1}$ while the crosses themselves remained stationary. Subjects were asked to report if the motion also appeared to spread outside the crosses. The results was that all of the twelve subjects reported it did.

**Experiment 6**

The spread of the motion outside the crosses in Experiment 5 brings with it a change in the shape of the aperture within which the motion is seen. The effect of aperture shape on motion direction is well known (Wallach, 1935). Diagonal stripes are seen to move orthogonally to their orientation when viewed through a circular aperture.
Figure 9. (a) Crosses filled with a texture of black and white diagonal stripes inserted into the gaps of the Ehrenstein figure. (b) The crosses of striped texture presented alone. (c) Crosses filled with a black and white dotted texture inserted in the gaps of the Ehrenstein figure. (d) The crosses of dotted texture presented alone. To observe the spreading effect, fixate the central cross in each case.
but when seen in a rectangular aperture, they appear to move in the direction of the longer side.

In this experiment, we examined the effect of context — crosses alone or embedded in the gaps of the Ehrenstein figure — on the perceived direction of motion of the texture in the crosses. The subjects were asked to report the dominant perceived direction or directions of the diagonal texture seen in Fig. 10a and b. When the crosses were inserted in the Ehrenstein figure, eleven of twelve subjects reported that diagonal motion was dominant. The remaining subject reported vertical and horizontal motions in the vertically and horizontally elongated parts of the crosses, respectively. When the crosses were presented alone, on the other hand, ten of twelve subjects saw only vertical and horizontal motion in vertically and horizontally elongated parts of the cross, respectively. The remaining two did not report any motion in the cross. The percentage of the subjects who saw diagonal motion for the textured crosses embedded in the Ehrenstein figures was significantly higher than for the textured crosses alone ($z = 3.79, p < 0.01$).

The change in the direction of perceived motion can be understood by considering the shape of the apparent aperture. When no spreading was seen for the crosses alone, the physical aperture imposed horizontal and vertical directions on the motion of the stripes. When the texture appeared to spread outside the physical cross, the apparent aperture became more circular and reduced its biasing effect on perceived direction of motion.

The perception of texture and motion transparency in Fig. 9 may reflect interpretations based on likelihood. In particular, Figs 9a and 10c could arise from two physically possible configurations. One is the exact alignment of textured crosses with a background grid. The more probable configuration (see Nakayama and Shimojo, 1990, for an analogous example) is that transparent disks overlie the intersections of a continuous regular grid. The disks are filled with white texture elements. In this configuration, the white elements on the transparent surface are invisible on the white background and only become visible when silhouetted against the black lines of the grid.
Figure 10. (a) Crosses filled with a texture of diagonal luminance gratings inserted into the gaps of the Ehrenstein figure. (b) The crosses of the luminance gratings presented alone. In both displays, the texture drifted but the cross-shaped aperture within which the texture was presented did not move.
GENERAL DISCUSSION

Three questions have been addressed: (1) How rapidly can the visual system establish the surface decomposition that accompanies subjective transparency? (2) Can this surface decomposition influence early levels of visual processing? (3) Can the visual system assign texture and motion attributes to subjective transparent surfaces?

The results from the pattern identification tests showed that identification was facilitated by transparency very quickly and that the facilitation depended at least in part on the presence of transparency. Transparency also influenced the McCollough effect, due presumably to the more effective continuity of horizontal and vertical stripes when seen as transparent overlays. This suggests that change of perceptual organization caused by transparency can affect very early levels of visual processing. Finally, we showed that the spread of surface qualities over a subjective surface could include factors of texture and motion as well as brightness and color.

Acknowledgement

This work was supported by AFOSR grant #91-0169 to P.C.

REFERENCES


Transparency and surface decomposition


