Three-dimensional spatial grouping affects estimates of the illuminant

Kenneth R. Perkins and James A. Schirillo

Department of Psychology, Wake Forest University, Winston-Salem, North Carolina 27109

Received July 16, 2003; accepted August 12, 2003

The brightnesses (i.e., perceived luminance) of surfaces within a three-dimensional scene are contingent on both the luminances and the spatial arrangement of the surfaces. Observers viewed a CRT through a haploscope that presented simulated achromatic surfaces in three dimensions. They set a test patch to be approximately three times more intense than a comparison patch to match the comparison patch in brightness, which is consistent with viewing a real scene with a simple lighting interpretation from which to estimate a different level of illumination in each depth plane. Randomly positioning each surface in either depth plane minimized any simple lighting interpretation, concomitantly reducing brightness differences to approximately 8.5%, although the immediate surroundings of the test and comparison patches continued to differ by a 5:1 luminance ratio.

The above brightness difference is consistent with what an observer would have perceived had the luminances of the surfaces varied in this way in a real three-dimensional scene if each depth plane contained a real illuminant. Thus we hypothesize that the increase in brightness difference that occurred in the above conditions was the result of enhancing the likelihood that these stimuli would contain a simple lighting interpretation from which to estimate a different level of illumination in each depth plane. We verified this hypothesis by randomly positioning each surface in the array in either of the two depth planes. In the real world, such reorganization would minimize a simple lighting interpretation. It had a comparable effect here in that it reduced brightness differences to a mere 8.5%, although the immediate surrounds of the test and comparison patches continued to differ by a 5:1 luminance ratio.

We claim that the small brightness difference remained because lower-level simultaneous contrast mechanisms emerged as the likelihood decreased that the scene would contain an unambiguous lighting interpretation from which to estimate different levels of illumination.
The current study uses luminance differences of nonretinal adjacent surfaces in three dimensions to extricate the local contributions to brightness\textsuperscript{12} evident in Gogel and Mershon's\textsuperscript{1} adjacency principle from Gilchrist's coplanar ratio hypothesis,\textsuperscript{16} which requires an estimate of the illuminant.\textsuperscript{11,13,14}

2. METHOD

A. Observers
Four observers with normal or corrected acuity (20/20) and normal stereo acuity were tested. Two paid Wake Forest University undergraduates (AL, a 22-yr-old female and DH, a 22-yr-old male) were inexperienced observers, and naive regarding the experimental design. The other observers were the authors (KP, a 22-yr-old male, and JS, a 41-yr-old male) and were knowledgeable about the experimental paradigm. JS had prior experience making brightness judgments on complex achromatic displays.

B. Apparatus
Stimuli were generated by a Power Macintosh 7600/132 and presented on an accurately calibrated Radius PressView 17SR 17-in. color monitor. The $832 \times 624$ pixel screen produced achromatic stimuli at CIE chromaticity $x = 0.27, y = 0.28$. The scan rate was 75 Hz noninterlaced. The chromaticity of each phosphor was measured spectroradiometrically. The red, green, and blue guns were linearized by use of an 8-bit lookup table. Luminance was approximately constant ($\pm 3\%$) within the central region of the screen that displayed the test and comparison patterns. Observers viewed the monitor through a haploscope at an effective viewing distance of 100 cm in a dark room. They could change the distance of two small 45-deg mirrors, one in front of each eye, to fuse the left- and right-eye images into a single crisp stereoscopic image.

C. Stimuli
The CRT simultaneously displayed four achromatic patterns on an otherwise dark screen. Identical $5^\circ \times 5^\circ$ patterns (composed of either squares or circles) were located in the upper-left and upper-right quadrants of the screen and contained 34 arc min of crossed retinal disparity. Two other identical $5^\circ \times 5^\circ$ patterns without retinal disparity were located in the lower left and lower right quadrants of the screen. The retinal disparity determined the extent that the upper array projected into a near depth plane compared with the perceived depth plane of the (far) lower array.

Figure 1(a) depicts the appearance of the fused uniform-luminance surface surround, Fig. 1(b) depicts the fused graded-luminance surface surround, and Fig. 1(c) depicts the fused randomly-mixed-luminance surround. In all three conditions the space-average luminance of the surfaces within a depth plane is constant. Figure 1(c) is the condition that most closely replicates the conditions in the study by Schirillo and Shevell.\textsuperscript{7} Used in the first experiment, it produced the maximum brightness differences that observers were able to obtain given two depth planes and constant local induction.\textsuperscript{11–14}

In experiments 1 and 2, each upper and lower array consisted of 25 $1^\circ \times 1^\circ$ diameter patches (either squares or circles). Experiments 3, 4, and 5 used 0.75$^\circ$ diameter circles. The simulated surfaces in each depth plane varied in luminance over a 6:1 range. In the graded and randomly mixed luminance conditions, each surface in the lower test array was set to one fifth (20%) of the luminance of its corresponding surface in the upper com-
comparison array. The resulting image, composed of lights covering a 30:1 luminance range, was sufficient for the two arrays to appear to have identical surface reflectances under different levels of illumination.\textsuperscript{17}

In all experiments, a 0.75° square patch was centered in both arrays. The observer manipulated the lower test patch, and the experimenter set the upper comparison patch to one of seven luminances, presented in pseudo-random order without replacement, using the method of constant stimuli. The comparison luminances were 16.44, 18.79, 21.14, 23.49, 25.84, 28.18, or 30.53 cd/m\textsuperscript{2}. All other surfaces of both arrays ranged in luminance from 7.44 cd/m\textsuperscript{2} to 31.42 cd/m\textsuperscript{2}. The geometric mean luminance across the entire display was 11.43 cd/m\textsuperscript{2}. The test patch and the comparison patch were always centered in the middle 1° × 1° region (referred to as the immediate surround) of their respective arrays. In all three surface conditions, the immediate surround of the upper comparison patch had a luminance of 14.55 cd/m\textsuperscript{2} and that of the lower test patch had a luminance of 2.91 cd/m\textsuperscript{2}.\textsuperscript{2}

D. Procedure

Observers did several practice sessions before beginning the reported measurements. They maintained a stable head position with a chin rest. Before each session they dark-adapted for 3 min and then light-adapted for 3 min to a homogeneous field at the luminance level of the geometric mean of the test- and comparison-surround luminances. Observers used a method of adjustment to vary the luminance of the lower test patch to match the upper comparison patch in brightness. They were told to adjust the test patch to appear identical to the comparison patch.\textsuperscript{7,18} They were also told to spend about the same amount of time looking at the upper and lower halves of the display by alternating their gaze between the two halves about once every 2 s. The naive observers were not told that the two halves of the display differed by a constant. The data show little variability, suggesting that adaptation was stable and complete.

Observers used a computer joystick to match the test-patch luminance in the lower (far) surround to the comparison-patch luminance in the upper (near) surround to a homogeneous field at the luminance level of the geometric mean of the test patch below the physical luminance of the comparison patch.\textsuperscript{17} This is how observers would have set a test patch in shadow that they perceived as brighter than a comparison patch under more illumination that they perceived as dimmer. Although a CRT simulation provides no actual illumination, these findings are consistent with observers having a simple interpretation of a hidden light source in the upper-right corner behind them in the graded condition [Fig. 1(b)], increasing the likelihood that they would estimate that each depth plane contained a different level of illumination.

Three observers judged the brightness of the test patch in the graded-luminance surround condition [Fig. 1(b)] as being roughly equivalent to that in the mixed-luminance surround condition [Fig. 1(c)]. [In Fig. 2, compare squares with diamonds: AL, \( p < 0.05 \); DH, \( p < 0.01 \); JS, \( p < 0.01 \); KP, \( p < 0.01 \).] Only the most experienced observer, JS, set the test luminance in the mixed-luminance surround condition significantly lower than in the other two surround conditions \( (JS, p < 0.01) \). Overall, these findings suggest that observers are more likely to estimate greater differences in the level of illumination in scenes containing simple lighting interpretations \( [i.e., \text{Fig. 1(b)} \text{ and 1(c)}] \), whereas when specific lighting cues decrease \( [\text{i.e., \text{Fig. 1(a)}],}\text{ the baseline effects of lower-level mechanisms (e.g., opponent processes) predominate.}^{1,12,13,19}\)

B. Experiment 2: Brightness Matching with a Circular Array

In experiment 1 squares were used to build on prior work done by Schirillo and Shevell.\textsuperscript{7} In experiment 2 a circle of comparable diameter replaced each square, minimizing adjacent borders to only four tangential points, so that when surface size is reduced in experiment 3, concomitantly eliminated border adjacency has minimal impact. Thus each 5° × 5° array now held 25 1° diameter circular patches. The diameter of each circle equaled the width of a square, reducing the total surface area to 79% of the squares. The black background of the screen was visible through the gaps between the circles. The 0.75° square test and comparison patches remained square, though both were centered in the middle 1° diameter circle (rather than square) of their respective arrays.
The results in these conditions (Fig. 3) were similar to those found in experiment 1, even though each circle touched its neighbors at only four points. The average percent deviation from a physical luminance match for the mean of the three conditions’ middle data point (i.e., 23.49 cd/m²) for all four observers was ~33%. The test-patch luminance in the uniform-luminance condition remained above that in the graded condition for three of the observers. (In Fig. 3, compare diamonds with squares: DH, $p < 0.02$; JS, $p < 0.01$; KP, $p < 0.01$; AL showed no significant difference: AL, $p > 0.15$). However, the test-patch luminance in the uniform condition was no longer set significantly above that in the mixed condition for any of the four observers. (In Fig. 3, compare diamonds with circles: AL, $p > 0.45$; DH, $p > 0.30$; JS, $p > 0.20$; KP, $p > 0.25$).
C. Experiment 3: Reducing Circle Size Reduces Brightness Differences

The diameter of the circles used in experiment 2 was reduced to 0.75°, leaving the total surface area of the smaller circles ~38% of the surface area of the larger circles. While each circle’s diameter was reduced, the spatial location of the center of each circle remained fixed. Thus the circles no longer abutted. The 0.75° square test and comparison patches remained centered within the immediate surrounds of the 1° diameter circle.

The reduction in total area, combined with the elimination of abutting points, substantially reduced the average percent deviation from a physical luminance match to ~22%. Moreover, there was no longer a significant difference between the test-patch luminance in the gradient condition and either of the other two luminance conditions with smaller circles. (In Fig. 4, compare squares with either circles or diamonds: AL, p > 0.50; DH, p > 0.46; JS, p > 0.19; KP, p > 0.20). As expected from related chromatic experiments and the adjacency principle, the frontoparallel separation of the circles reduced local contrast, thereby reducing brightness differences.

D. Experiment 4: Ungrouping Coplanar Surfaces Reduces Brightness Differences

The previous three experiments retained some probability of a simple lighting interpretation by spatially grouping the upper and lower arrays into near and far depth planes, respectively. Experiment 4 minimized such an interpretation by randomly assigning each small circle to a depth plane by using 34 arc min of retinal disparity, so that over the entire field, half of the circles appeared in the near plane and the other half appeared in the far plane (Fig. 5). The test patch and its immediate surround and the comparison patch and its immediate surround remained in their respective far and near depth planes.

Surrounds of either mixed or uniform luminance in this stimulus configuration produced brightness matches that deviated from a physical luminance match by a mere 10% (Fig. 6, open symbols). This finding is consistent with creating stimuli that have highly improbable coplanar lighting interpretations, so that their smaller brightness differences reflect primarily processes driven by local contrast. It is important to realize that these reduced brightness differences occurred despite the fact that the immediate surround of the test and comparison patches remained at a 5:1 luminance ratio. That is, local contrast was the same in this experiment as in the earlier three.

Interestingly, greater brightness differences occurred for the two experienced observers in the graded-luminance surround condition. (In Fig. 6, compare squares with either type of open symbol: JS, p < 0.01; KP, p < 0.01). This result suggests that the enhanced likelihood of an illumination difference between depth planes derived from a cue that an illumination source is off to the side in the graded-luminance condition may be experience dependent (AL, p > 0.24; DH, p > 0.08).

E. Experiment 5: Enhanced Brightness Differences with Monocular Viewing

To determine the extent that randomly arranging surfaces of different luminances in stereo depth reduced the likelihood of a simple lighting interpretation, observers in experiment 5 monocularly viewed the same stimuli used in experiment 4. They viewed only one half of the CRT screen through the haploscope by wearing a patch over
the other eye, which eliminated stereo depth. Two of the observers viewed with the left eye, and the other two viewed with the right eye.

Compared with experiment 4, eliminating stereo depth made the test patch appear brighter or the comparison patch appear dimmer or both (Fig. 7, all symbols). That is, the test-patch luminance in all three luminance profiles was set below the physical luminance match by \(-22\%\) on average. This suggests that observers are likely to interpret a two-dimensional scene with this configuration of surface luminances as containing more than one illuminant. Strengthening this hypothesis is the fact that observers set the test-patch luminance slightly darker in the graded condition, suggesting an additional shift in their estimation that the test and comparison patches are under different levels of illumination. (In Fig. 7, compare squares with open symbols). Interestingly, simply adding a third dimension can actually make this lighting interpretation less likely. That is, such a scene would now contain a collection of surface luminances in depth that would make it highly improbable that there would be two levels of illumination, one for each depth plane.

4. DISCUSSION

The above findings indicate that brightness matches between test and comparison patches covary with changes in the luminances of surrounding surfaces in a way that is consistent with what a corresponding shift in an estimate of the illuminant would predict. That is, the brightness measurements covaried in the same way as they would in scenes containing different real illuminants.

We found that brightness matches between a test patch in one depth plane and a comparison patch in another depth plane, with a constant 5:1 luminance ratio across depth planes (experiment 1, Fig. 1), were consistent with what observers would have produced had they estimated a different level of illumination in each depth plane. This effect was largest when the graded luminances of the surfaces within a depth plane were randomized [Fig. 1(c), adapted from Ref. 7]. Subsequent experiments demonstrate the extent to which different spatial luminance arrangements reduce the likelihood of these estimated differences in the level of illumination while maintaining constant local induction. This suggests that bright-
ness judgments are the result of multiple mechanisms: a local opponent process and a more global process that includes estimating the level of illumination.

Many psychophysical studies use CRTs rather than real papers and lights to simulate surfaces grouped by similar luminance, without any actual illuminant, to create the illusion of regions that differ in illumination. In the current study, using a CRT offered more stimulus control than would have been possible with paper. However, higher-order processes may limit the degree to which observers accept that a CRT simulation has multiple levels of illumination. Thus, although the current findings are consistent with how brightness differences would increase with an increase in a real difference in the level of illumination, they may underestimate the magnitude of brightness shifts that would be possible in real scenes as observers’ estimates of different levels of illumination increase.

Fig. 7. Brightness matches of 0.75°-diameter circular arrays for mixed (circles), graded (squares), and uniform (diamonds) surrounds randomly assigned to different depth planes under monocular viewing. Observers: AL, DH, JS, and KP.

Fig. 8. Only the mixed-luminance-condition brightness matches with 1° diameter circular patches (circles), 0.75° diameter circular patches (squares), and 0.75° diameter circular patches in randomly assigned depth planes (diamonds). Observers: AL, DH, JS, and KP.
Although it is impossible to ensure that observers judged the brightnesses and not the lightnesses of the surfaces in question, the current results are comparable with brightness matches but not lightness matches made in prior related work, where independent judgments of both surface qualities were made. That is, when a scene promoted a simple lighting interpretation that each depth plane contained a different level of illumination, lightness matches approximated reflectance matches. Unfortunately, the present design cannot reaffirm these notions, since collecting lightness measurements would require informing observers of differences in illumination. Yet recent evidence by Rutherford and Brainard, relating lightness judgments to direct measures of perceived illumination, dovetails nicely with our current hypothesis.

Replotting only the mixed-luminance surround conditions across the three main experiments reiterates the significance of our main findings (Fig. 8). The large-circle display produced brightness differences of ~33% (Fig. 8, circles), replicating Schirillo and Shevell results. This represents what may be the maximum extent that illumination could have been estimated to differ across the depth planes in our displays. As expected, reducing the circle size reduced local contrast and thereby brightness differences to ~25% (Fig. 8, squares), since the total surface area of each circle decreased, and the circles no longer abutted. Randomizing the spatial arrangement of the circles’ mixed-luminance distribution in depth reduced brightness differences further to a mere 8.5% (Fig. 8, diamonds), which is consistent with the notion that randomizing retinally nonadjacent circles minimizes the ability to estimate different levels of illumination across a three-dimensional image. This becomes evident when viewing the display with one eye covered restores brightness differences to ~17%.

The current stimuli are unique in that darkness surrounded each circle, which suggests that manipulating the estimate of the illuminant need not rely solely on edge integration across the scene. Recent physiological evidence supports the idea that retinally nonadjacent (i.e., noncontiguous) luminance signals combine in visual cortex, where estimates of illumination probably occur. Thus a plausible model of the above findings would be that local opponent processes produce a relatively small, maybe fixed, amount of possible brightness differences, which can be augmented as the likelihood increases that different levels of illumination are falling on the surfaces being judged.

ACKNOWLEDGMENTS

Special thanks to Steve Shevell for initially suggesting this paradigm. Thanks to David Brainard, Larry Maloney, and Steve Shevell for helpful insights and comments on earlier versions of the manuscript and to the reviewers for their thoughtful observations. Thanks also to Andrea Lindsay and David Hall for serving as observers and to Guy Buss for computer programming.

Address correspondence to James A. Schirillo at the address on the title page; phone, 336-758-4233; fax, 336-758-4733; e-mail, schirija@wfu.edu.

REFERENCES