The interaction of color and luminance in stereoscopic vision

Cary Lu and Derek H. Fender

In this paper, two major areas of vision research traditionally examined in isolation are combined. The interaction between binocular vision and color vision is mapped to determine the way in which visual information based on differences in color and visual information based on differences in luminance are used to perceive depth.

There are three major physiologic cues to depth perception: the accommodation of the eyes, convergence of the eyes, and retinal disparity. Of these, accommodation can be a monocular cue, the others are binocular. Accommodation and convergence are not essential factors in seeing depth, but retinal disparity is important. If an object viewed with both eyes is perceived as a single object, we can derive the horopter connecting all objects which, for the same convergence, can be seen as fused. The horopter contains those objects in visual space that are imaged without retinal disparity.

Single vision is also subserved by the ability of the central visual system to fuse objects that are not exactly on the horopter. Retinal correspondence need not be exact for single vision, only close enough for there to be correlation between the two images. Point-by-point remapping of these

From the California Institute of Technology, Pasadena, Calif.

Supported in part by Grants CM-01335 and NB 03627 from the National Institutes of Health, United States Public Health Service.
differences in disparity by the central visual system would yield the cue for depth perception. We can postulate a rather simple remapping mechanism which would operate in the case of simple figures having much monocular form structure, such as the classical Helmholtz line drawings of stereo pairs. These drawings typically show a pyramid seen from above. Once the outlines of the bases of the pyramids in the two images are fused, depth information concerning the remainder of the pyramidal structure could be computed by noting the horizontal distance between each closely spaced pair of lines. However, the experiments of Julesz with random-dot stereo-pair patterns have shown that the process of extracting depth information in the human visual system must be much more complex than the previous speculation. With random-dot stereo patterns it is not sufficient merely to look for local displacements left or right for coincidence between dots, for such coincidences can occur randomly in either direction and over many increments of distance. Instead, the depth sense extracts a global solution rather than a local one. It detects that complete correspondence between certain areas can be established if whole areas are shifted laterally even if this conflicts with the point-by-point solution. Whether or not these two processes are actually operative in the visual system is still an undecided question and whether these processes are really distinct is doubtful. However, it is quite clear that in some respects the two processes are different. In the first case the contours which delineate form can be seen monocularly, whereas in the case of the Julesz patterns the contours which are generated by the depth percept cannot be seen monocularly.

In real life there are obviously many monocular contours available, and thus a large part of depth perception may depend on a process similar to the first described above. Conversely, there are cases that are much closer to the stimulus presented by the Julesz patterns. For example, viewing the leaves of a tree at a distance: In this case, the individual leaves and their contours cannot be matched individually between the left and right eye images, for within the limits of acuity one leaf may very well match several others with equal goodness of fit, and yet the three-dimensional structure of the tree can be seen. In this case, the matching procedure is more likely to be akin to the global process of the Julesz type of patterns. We may therefore consider which parameters of the visual scene are used in the central matching process to fit together the two retinal images in a case such as this. If we refer to this as a correlation process, then the correlation between the two images may be performed on the luminance distribution of the two patterns only or the color distribution of the two patterns, or there may be a trade-off between these parameters.

It is the purpose of this paper to explore the interaction of these parameters in the formation of a depth percept with the use of Julesz type of patterns, in which depth cannot be inferred from monocular information.

Method.

Random-dot stereo-pair patterns. The patterns used for the experiments reported here consisted of two figures, each with a grid of one hundred by one hundred squares which were randomly white or black with equal probability. The right eye pattern was the same as the left eye pattern, except that a 40 x 40 area at the center of the pattern was shifted four squares to the left.

The patterns used for these experiments were produced by computer-generated microfilm with the use of an Information International FR-80 which makes direct photographs of a CRT face. A complementary or negative pattern was also used in the experiment. This was the same as the original pattern except that white squares were substituted for the black squares of the original, and vice versa. The complementary pattern was produced by inverting arrays in the computer, not by negative-positive photographic techniques, to avoid spread of the image by-halation and other effects.

Target presentation. Four patterns were used, two to generate the image for each eye. For each side the patterns were complements of each other, thus each side produced a uniform field if viewed with the same light source. Between sides, the patterns constituted a stereo pair.

Optical system. The complete optical system is
shown in Fig. 1. In a typical experiment, the left patterns, L1 and L2, were lined up with micrometer adjustments, using the same light source, until the subject reported a field uniform in appearance. The same was done for the right patterns, R1 and R2. The left and right sides were then adjusted so that the images converged optically at infinity. Filters for control of color and luminance were inserted between M1 and H1 and between similar positions on the right side.

For all experiments, the field of view of the patterns was ten degrees. The optical distance of the patterns was 1 M. The field outside the patterns was not illuminated except for minor scattered light, which was less than two per cent of the illuminated portions of the patterns.

**Experimental procedures.** For each observation, the luminances of L1 and R1 were maintained at a constant value; the subject adjusted the luminances of L2 and R2 by servo-driven neutral density wedges at W and at a similar position on the right side. The two servo-controllers were driven in step by electrical interlock. In general, the subject observed that as the luminances were adjusted depth was easily perceived when pattern 2 was dimmer than pattern 1. As the luminance of pattern 2 was increased, a point was reached at which the depth could no longer be perceived. Only when pattern 2 was at a higher luminance than pattern 1 was the percept of depth regained.

For each particular experiment, pattern 1 was set to the highest luminance possible; the subject then adjusted pattern 2 from a matching luminance through increasing values until depth could just be perceived.

Experiments began with high luminance in pattern 1, and after an adjustment of pattern 2 and measurement of luminances, pattern 1 was reduced in luminance and pattern 2 readjusted, always from a matching luminance setting. This process was repeated until the patterns were too dim to be easily seen, either monocularly or binocularly. Typically, the luminance of pattern 1 would begin at about 250 footlamberts and end at about 10 footlamberts. About 15 minutes was required for a series of these measures with a given set of colors.

There were two possible threshold criteria: the point where depth could just be seen, and the point where depth could just not be seen. In all the data reported here, the measurements are for the case where depth could be seen.

The criterion for seeing depth was subjective; however, there was a simple method for providing a direct comparison between the test stimulus and a stimulus that did have easily perceived depth. This involved simply turning off the source light to either pattern 1 or 2, so that the images were then seen as some color and luminance against black. This always resulted in easy depth perception. When the blanked pattern was turned on again, the subject judged if the percept of depth remained. In practice, it was found that subjects were fairly consistent in their criterion for determining the threshold for depth perception.

**Subjects.** All subjects were experienced observers familiar with the operation of the experimental apparatus, but most were not familiar with the purpose of the experiment. All subjects had no difficulty or ambiguity in identifying depth using black and white Julesz patterns. Quantitative data were recorded from five subjects; qualitative data from an additional six subjects were used to confirm the basic finding: depth perception with
Results

Table I displays the results of a typical experiment in which red (634 nm.) and green (512 nm.) were used. It will be noticed that the mean contrast ratio for the percept of depth is significantly less when the green areas of the pattern are brighter than the red areas compared with the opposite configuration. Table I also shows that there is no major relationship between luminance and the required contrast ratio for stereopsis; hence in the following work the mean contrast ratio over all luminance values will be used as representative of the given color-pair combination.

The contrast ratio for red-brighter-than-green and the corresponding value for green-brighter-than-red can be regarded as points on a continuous scale of contrast ratios (Fig. 2). In the middle of the scale, the contrast ratio is unity and red and green images are of equal luminance. There are three regions on the scale: The center range corresponds to contrast ratios for which depth is not perceived, and the two regions on either side are values for which depth is perceived.

The experiment described above can be repeated for red versus many other colors (call these colors C1 and the fixed color red, C2). A series of contrast ratio scales of the type shown in Fig. 2 can be plotted vertically against the wavelength of C1 as in Fig. 3. Above the horizontal axis (not drawn in the diagram), the contrast ratios are plotted for patterns with the red image (C2) brighter than the image of color C1 on the horizontal axis; below the axis, the

### Table I. Luminances of pattern 1 and pattern 2 at the threshold for depth perception

<table>
<thead>
<tr>
<th>Luminance (footlamberts)</th>
<th>Contrast ratio pattern 1/pattern 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Pattern 1 Pattern 2</td>
<td></td>
</tr>
<tr>
<td>(red) (green)</td>
<td></td>
</tr>
<tr>
<td>256.47 26.40</td>
<td>9.78</td>
</tr>
<tr>
<td>167.25 13.00</td>
<td>12.90</td>
</tr>
<tr>
<td>104.25 8.95</td>
<td>11.68</td>
</tr>
<tr>
<td>54.40 4.62</td>
<td>11.75</td>
</tr>
<tr>
<td>34.21 3.86</td>
<td>8.80</td>
</tr>
<tr>
<td>22.12 2.42</td>
<td>9.47</td>
</tr>
<tr>
<td>13.08 1.42</td>
<td>9.17</td>
</tr>
<tr>
<td>Mean contrast ratio</td>
<td>10.51</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.60</td>
</tr>
<tr>
<td>B. Pattern 1 Pattern 2</td>
<td></td>
</tr>
<tr>
<td>(green) (red)</td>
<td></td>
</tr>
<tr>
<td>199.44 151.16</td>
<td>1.32</td>
</tr>
<tr>
<td>97.36 45.38</td>
<td>2.17</td>
</tr>
<tr>
<td>69.00 45.08</td>
<td>1.53</td>
</tr>
<tr>
<td>45.03 31.33</td>
<td>1.45</td>
</tr>
<tr>
<td>30.20 22.15</td>
<td>1.37</td>
</tr>
<tr>
<td>19.30 15.03</td>
<td>1.27</td>
</tr>
<tr>
<td>11.91 7.89</td>
<td>1.48</td>
</tr>
<tr>
<td>Mean contrast ratio</td>
<td>1.51</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Red = 634 nm. and green = 512 nm. Subject C. L.

A: Red set brighter than green, B: Green set brighter than red.
color C1 of the image represented on the horizontal axis is brighter than the red image (C2). There are now three areas in the graph. Between the plotted lines, depth is not perceived; above the upper line and below the lower line, depth is perceived.

Similar data are available for two other values of the color C2, green (512 nm.), and blue (477 nm.).

If for one subject the three data sets for which C2 is red, green, and blue, respectively, are put together, a three-dimensional cartesian plot can be obtained in which two axes represent the wavelengths of the colors C1 and C2 and the third axis is the contrast ratio. The result is shown in Fig. 4 for the case of the image of color C2 brighter than the image of color C1.

If the values of contrast ratios below the axes of diagrams such as Fig. 3 are used, another surface can be constructed as in Fig. 5. This surface should be mirror antisymmetric to the surface of Fig. 4 about the line $\lambda (C_1) = \lambda (C_2)$ where $\lambda (C_n)$ means "the dominant wavelength of color $C_n". This lower surface plots the case of the image of color C1 brighter than the image of color C2, therefore the position of data point $\lambda (C_1) = 634$ nm., $\lambda (C_2) = 477$ nm. above the base plane in Fig. 4 is equivalent to the position of $\lambda (C_1) = 477$ nm., $\lambda (C_2) = 634$ nm. below the plane in Fig. 5. This theoretical requirement of symmetry is reflected in the experimental data within the limits of reliability of the observations. The surfaces plotted in Fig. 4 and Fig. 5 separate the condition of depth-not-perceivable between the surface and the base plane from all other regions of the data space in which depth is perceivable.

A section through the surface for $\lambda (C_1)$
Color and luminance in stereoscopic vision

Fig. 4. Contrast ratio threshold for stereopsis. Red, green, and blue on axis C₂ versus color of wavelength shown on axis C₁. Image of color C₂ brighter than image of color C₁.

\[ \lambda (C_2) \] is shown for two subjects in Fig. 6. Along this section, the color used in all channels is the same; hence only luminance differences are involved. The contrast ratios are about the same as those found with the use of light from a tungsten filament, about 2,800 degrees Kelvin correlated color temperature, for all channels. The variation of contrast ratios with color shown in this plot is small and is not significant.

Discussion

The result that depth perception is possible with Julesz patterns only when the individual elements are delineated monocularly by a luminance contrast and not from monocular contours created by color differences alone is unexpected from the psychophysical literature. It has always been tacitly assumed that the detection of contours is just a matter of varying colors and luminances for detectable borders. While this may be valid for monocular contours, stereo contour detection is now shown to be quite a different matter.

The evidence presented here suggests independent coding for color and luminance in the visual system and that binocular vision uses the luminance channel principally with secondary effects based on the particular colors involved.

The neurophysiologic literature contains a certain amount of indirect evidence in support of this conclusion. At the retina of several animals, including the frog, snake, and guinea pig, Granit found that electrophysiologic recordings supported the hypothesis of two types of fibers: the dominators which had broad sensitivity throughout the spectrum, corresponding to luminosity curves, and modulators which had narrower responses to restricted portions of the spectrum. In the lateral geniculate, DeValois and associates found two general types of cells: opponent color cells with differential responses to varying spectral stimulation, and nonopponent cells with the same responses irrespectively of the spectral stimulation.

Hubel and Wiesel found that most of the cells in the primary visual cortex of the macaque and spider monkey responded to
spatially appropriate stimuli without regard to the color of the stimulus. They estimate that only about seven per cent of the cells respond to color. Roughly half of these color-responsive cells had responses to a specific color similar to their response to white light. The other half responded more selectively, such as only firing to moving bars of a particular color and not to white light.

The conclusion that color information regarding simple contours is discarded before binocular contour information is processed may apply only to the particular binocular stimulus configuration without monocular cues used in these experiments. It certainly does not apply to all binocular stimuli which contain monocular cues; indeed, Helmholtz stereograms in two colors readily generate the percept of depth when matched for luminance.

The same-color plots shown in Fig. 6 are similar to plots of the luminance difference or contrast ratio required to see monocular borders in a bipartite field of uniform spectral color. The contrast ratios obtained in the present experiments are larger than the contrast ratios for monocular experiments.

Binocular contours are based on monocular information, hence the higher threshold is surprising since a better signal-to-noise ratio should be available in binocular viewing. However, a higher threshold is consistent with the hypothesis that color information is not the primary cue for stereo contours.

The same-color contrast threshold is somewhat larger than that for some of the combinations of different colors. This raises the possibility that some color difference information is used in depth perception, enough to lower the threshold contrast ratio under certain conditions. However, it is interesting to note the way in which this interaction occurs; if the wavelength separation is large (red against blue), with the longer wavelength image brighter, then the color information does not supplement the luminance signal and a large contrast ratio is required. As the wavelength separation is reduced, some supplementation occurs and lower contrast ratios are adequate for stereopsis. If the shorter wavelength image is brighter, then supplementation takes place for all wavelength separations.
Color and luminance in stereoscopic vision

Fig. 6. Section of Fig. 4 for which $\lambda (C_i) = \lambda (C_j)$; that is, no color differences, luminance difference only. Solid line, Subject C. L.; Chain line, Subject J. R. C.

Fig. 6. Section of Fig. 4 for which $\lambda (C_i) = \lambda (C_j)$; that is, no color differences, luminance difference only. Solid line, Subject C. L.; Chain line, Subject J. R. C.

present there is no satisfactory explanation for this asymmetry of response.

Although the existing psychophysical literature does not provide information about the interactive effects of color and luminance in binocular vision, there is some information about the possibility of independent processing of color and luminance information in other visual tasks. The original psychophysical opponent color theory by Hering, developed by Hurvich and Jameson, is based on the independence of color and luminance channels but does not specify the extent of such independence. Several studies of chromatic adaption have suggested independence of color and luminance systems.

All of these psychophysical data lend support to the hypothesis of independent color and luminance channels in the visual system, although none touch on the question of binocular perception when all monocular portions of the image are perceived intact.

From the evolutionary standpoint, the perception of luminance is more fundamental than color, and in the evolution of mammalian visual systems, binocular vision apparently occurs lower than color. These results suggest that color was not added to stereo contour detection but was added to simple or monocular contour detection.

Given the information-handling capacity of the visual system, there is an argument of economy that color information be disregarded in depth perception. Depth perception requires very fine analyses of retinal disparity to localize an object. If color is used, there must be at least two additional informational routes in the neural substrates of depth perception. This implies an increase in the complexity of the neural wiring with little useful information to be gained, since the color of an object is already identified by simple contour detection. Hence much simpler neural wiring can be used for nearly all visual situations with little penalty, and as the results here indicate that color and relative luminance are used to identify objects, luminance alone is used as the principal signal to determine their depth.

REFERENCES

Mapping the visual sensory onto the visual motor system

Gerald Westheimer and Sidney M. Blair

Recording in the brainstem of the alert macaque revealed single units whose impulse rates vary with eye position with preferred directions that are neither horizontal or vertical nor do they correspond to the direction of activity of any of the twelve extraocular muscles. The facts that these units change firing rates well before the occurrence of changes of eye position and that they are located cranial to the third nerve nucleus suggest that they are supranuclear. It is shown how these units could function in the processes by which the sensory visual input is translated into the pattern of oculomotor organization demanded by the kinematic laws of eye movement.

In presenting some of our work on single units in the mammalian oculomotor system, we shall take our point of departure from a fundamental characteristic of the oculomotor system: the manner by which a steady-state eye position is achieved in the alert primate. It is doubtful whether this is a reasonable model of all motor control, and from perusal of the literature on muscular control one wonders whether it is worth the trouble to make any comparison at all between the general somatic motor apparatus and the eye musculature. At any rate, on the basis of some behavioral experiments in the human being and the then state of knowledge of physiology, one of us concluded twenty years ago that to each steady-state eye position there corresponded a graded set of steady-state innervations to the eye muscles. This was confirmed in three independent papers published in 1970. All showed that in the alert macaque the impulse rate in motoneurons leading to eye muscles is a unique and even almost linear function of eye position. The most usual kind of eye movement, the saccade, involves the simultaneous resetting of impulse rates in all concerned motoneurons to a new value. Because the impulse rates uniquely code for eye position, one may talk about the position-coded output of ocular motoneurons. In view of the con-