Pergamon

0042-6989(95)00056-9

Vision Res. Vol. 36, No. 1, pp. 53–66, 1996

0042-6989/96 $9.50 + 0.00

Localizing Contours Defined by More Than One Attribute

JOSÉE RIVEST,*† PATRICK CAVANAGH‡

Received 14 September 1994; in revised form 23 February 1995

Two experiments were run in order to test how information from different attributes is combined to localize contours. In Expt 1 the apparent position of a contour defined by one attribute was measured while a contour defined by another attribute was presented beside it. Interactions were found between all pairings of luminance, color, motion and texture. These results suggested that the information associated with each contour is integrated at a common site. In Expt 2 the precision of localization was measured for contours defined by one, two or three attributes (combinations of luminance, color and texture). The improvement in precision with additional attributes again supported an integration of contour information at a common site prior to a decision of localization.

Localization Precision Spatial interaction Attributes

INTRODUCTION

In natural images, it is easy to distinguish one object from another because each object has different surface attributes such as color, luminance or motion. The visual processing of such attributes may be separate (e.g. Livingstone & Hubel, 1988; Maunsell & Newsome, 1987; Schiller & Colby, 1983; van Essen & Maunsell, 1983). For example, Maunsell and Newsome (1987) suggest that the parvocellular stream (LGN to V4) is specialized for form and color and the magnocellular stream (LGN to MT) may be specialized for motion. Moreover, specific losses of vision occasionally follow brain lesions. Patients have shown independent losses of vision for motion (Botez, 1975; Zihl, von Cramon & Mai, 1983), color (e.g. Damasio, Yamada, Damasio, Corbett & McKee, 1980; Mollon, Newcombe, Polden & Ratcliff, 1980; Pearlman, Birch & Meadows, 1979), and luminance (Rovamo, Hyvärinen & Hari, 1982). Regan, Giaschi, Sharpe and Hong (1992) found some patients with parietotemporal lesions who experience difficulties recognizing motion-defined letters but who have no problem recognizing luminance-defined letters.

In order to achieve reliable visual analyses, it could be advantageous to combine information from these separate analyses at a common site. Some studies suggest that information about different attributes is indeed combined. Yeh, Chen, De Valois and De Valois (1992), for example, examined the apparent positional shift induced by one adaptation blob on a subsequently presented test blob. In one condition, the adaptation and the test blobs were defined by color and luminance respectively, whereas in the other condition, the relations were reversed. They reported a positional shift in both of these inter-attribute conditions confirming that signals from color and luminance do interact in determining spatial position. Landy (1993) also showed that signals from texture and luminance are combined. He found that moving a contour defined by one attribute relative to a contour defined by another attribute shifts the perceived location of the contours.

Of particular interest for this paper are the possibilities that final decision of contour localization originates from a common site where information coming from different visual attributes is united, and that the precision of localization is improved as a consequence of this combination. Experiment 1 examines whether the position of one contour defined by one attribute (e.g. color) influences the position of a nearby contour defined by another attribute (e.g. luminance). Experiment 2 examines whether precision to localize a contour is improved as additional attributes are superimposed to define the contour.

EXPERIMENT 1

Experiment 1 examined whether the position of a test contour defined by one attribute (e.g. luminance) is influenced by the position of a flanking contour defined by another attribute (e.g. color). The contours were defined by luminance, color, motion or texture. The perturbation technique used by Badcock and Westheimer (1985) and Rentschler, Hilz and Grimm (1975) to

*Department of Psychology, Glendon College, York University, 2275 Bayview Avenue, Toronto, Ontario, Canada, M4N 3M6.
†To whom all correspondence should be addressed.
‡Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, U.S.A.

test spatial interaction was adapted to study whether attributes are combined together for spatial localization. In the original perturbation technique, two contours defined by luminance were presented very close to one another. Observers must judge the position of one of the contours—the test contour—while the other contour—the flanking contour—was presented at different distances from it.

Badcock and Westheimer (1985) showed that when the distance between the two luminance-contours is small (i.e. <3.0–4.0 min arc), the test contour appears closer to the flanking contour (attraction). When the distance between the contours is larger, the test contour appears further away from the flanking contour (repulsion). Rentschler *et al.* (1975) showed attraction for line separations of <10.0 min arc. For larger separations of the lines, they showed repulsion for one observer out of two. Similar displacement between stimuli presented at close proximity was also reported much earlier using the technique of figural aftereffects (e.g. Day, 1962; Ganz & Day, 1965; Gibson, 1933).

Attraction between two closely-presented stimuli has often been assumed to result from a linear summation of the profile of activity of neurons (e.g. Badcock & Westheimer, 1985; Hines, 1976; Hock & Eastman, 1995; Tyler & Nakayama, 1984). This interaction between the profiles of cortical activities produced by two contours is schematically illustrated in Fig. 1. On the two graphs to the left, the profiles of activity of two contours are represented separately. On the graph to the right, both contours are represented together at a common neural site; the profile of activity is the result of a linear summation of the two separate profiles of activity. For each graph, the x-axis represents different positions on the surface of the cortex and the y-axis represents the amount of activity at a given position on that cortical surface. It is assumed that localization decision depends on the position of the peak of the profile of activity (the mean or mode of the distribution). Notice that in the two graphs to the left, each peak signals a different position: the peak signalling contour 1 is at a certain distance to the left of the one signalling contour 2, but in the graph to the right where both separate profiles of activity have been summed, the two peaks are closer from each other due to the linear summation. After linear summation, the first peak in the summed profile shifts towards the second peak in proportion to the slope of the second profile at the location of the first peak (e.g. Tyler & Nakayama, 1984). [The mechanisms involved in repulsion may not depend on the summation of the profile of activity (see Badcock & Westheimer, 1985; Hock & Eastman, 1995 for details).]

Since the proposed spatial interaction between two contours depends on the summation of their respective profiles of activity, the spatial scale of the interaction will vary with the width of the profiles of activity. In support of this property, Banton and Levi (1993) showed inter-contour interactions within a larger range of offsets

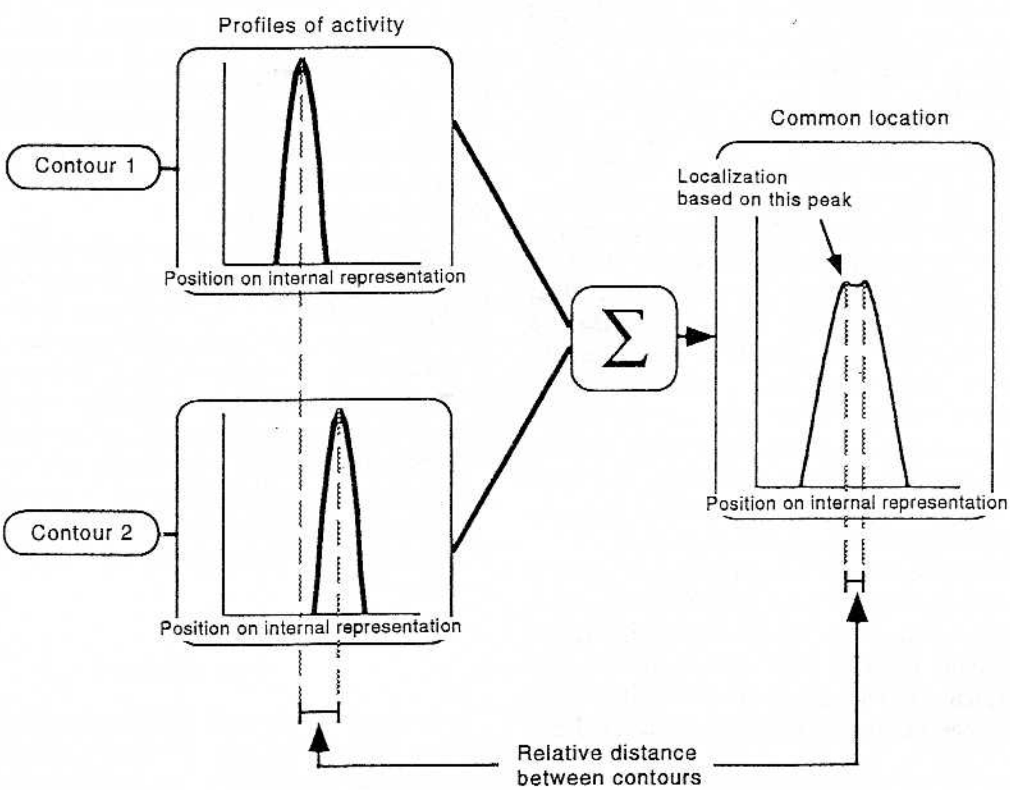

FIGURE 1. Illustration of a perceptual shift (attraction) between two contours defined by different attributes. Notice that the peaks of the separate profiles of activity of contours 1 and 2 are farther apart (graphs to the left) than when the two profiles are summed (graph to the right).

between motion-defined bars (14.0 min arc) than between luminance-defined bars (7.0 min arc). Receptive fields of directionally selective neurons are typically larger than nondirectionally selective neurons (e.g. Albright & Desimone, 1987; Anderson & Burr, 1987).

In Expt 1, the apparent position of a test contour defined by one attribute was measured while the position of an adjacent contour defined by another attribute was varied. Interaction between the position of contours defined by luminance, color, motion and texture was tested. If the profiles of activity associated with contours defined by different attributes are summed at a common neural site, displacement between the contours will be found. Moreover, if luminance really determines localization of contours as suggested by Gregory (1977) and Gregory and Heard (1979), a contour defined by another attribute (say color) should be perceived at the location of the luminance-contour. Conversely, a contour defined by color should have no influence on the perceived position of the luminance-contour. To avoid biases which might favor attributes according to their quality in the image, each attribute was presented at a contrast that produced about the same precision of localization. Consequently, whether luminance plays a privileged role in localization of contours was evaluated when its advantage of spatial resolution at high contrast was removed.

A control condition, in which the test and the flanking contours were both defined by luminance was run with one observer. This condition was run in order to see whether it was possible to replicate the interaction between two luminance-contours—as found by Badcock and Westheimer (1985) and Rentschler *et al.* (1975)—using our display.

Method

Subjects

Three members of the Department of Psychology at Harvard University were tested (JR, TWB, and ZH). They had normal or corrected to normal acuity and normal color vision. Observers TWB and ZH were naive concerning the purposes of the experiment. Observer JR was one of the authors. Only JR was tested in the control condition.

Apparatus

A Datacube image processor run by a Macintosh II cx was used. The experimental display (130.0 min arc height by 138.0 min arc width) was presented on a 19 in. Mitsubishi (Diamond Scan) color monitor. Observers were seated with their head and chin supported by a rest.

Except in the control condition, the top half of the display was filled with dark and light dynamic dot texture. This texture consisted of square dots (2.1 min arc on a side) randomly chosen to be dark or light (half dark, half light). The texture was replaced every 50.0 msec so that it was twinkling randomly. The contrast between the dark and light dots was always 50%. The mean luminance of the texture was 40.0 cd/m^2.

The properties of the texture on the right side of the display were always as previously described, but the properties of the texture on the left side were varied such that two adjacent contours were created. For example, in Fig. 2(a), a luminance test contour was created by making the mean luminance of the dots darker on the left side. Moreover, a color flanking contour was created by making part of the darker side greener while keeping it the same luminance.

Precision was made about equal among all attributes. First, each observer adjusted the luminance contrast 10 times until they could just see the contour, and the precision to localize this low-luminance contour was measured. Second, the color, texture, and motion "contrasts" that gave approximately equal precision of localization were found. The "contrasts" for all attributes are respectively defined in the following paragraphs.

Luminance. The "luminance contrast" creating a luminance-contour was a decrease in the mean luminance of the texture on the left side. The luminance contrast between the right and left regions was decreased minimally such that the test contour was set just above detection threshold for each observer. All observers needed a luminance contrast of 20%. This "luminance contrast" is larger than the contrast usually needed for detection because the luminance contour was presented within a dynamic noise.

Color. The "color contrast" creating a color-contour was a change in the green saturation on the left side of the display. The contour was created by making the left side unsaturated green. The green saturations used were between 16% and 21%; 0% being white (CIE coordinates: $x = 0.332$, $y = 0.333$) and 100%, arbitrarily defined as the chromaticity of the green phosphor alone (CIE coordinates: $x = 0.284$, $y = 0.578$). For each observer, the relative luminance between the gray and green sides was adjusted to maintain equiluminance at all saturations (color contrast). Each observer set the relative luminance between the green and gray sides of the display such that the contour was minimally visible. This adjustment was done when the green was at about 16% green saturation and the dot contrast was at 50% and ensured that the green side appeared equiluminant to the gray side for each observer.

Texture. For all observers, the "texture contrast" creating a texture-contour was a 400% increase in the height of the dots on the left side (height of 9.2 min arc and width of 2.1 min arc) compared to those on the right side while all other properties of the two textures were equal.

Motion. The "motion contrast" creating a motion contour was a change in the texture from twinkling to moving coherently upward or downward at a specific speed. The "motion contrast" is defined by the speed of the coherent motion. The speeds used were between 5.0 and 13.0 min arc/sec. The direction of motion was reversed at each trial in order to avoid motion aftereffects. Even though motion is the most obvious cue differentiating the two half-fields in this stimulus, a time-averaged (time integrated) representation of the two half-fields

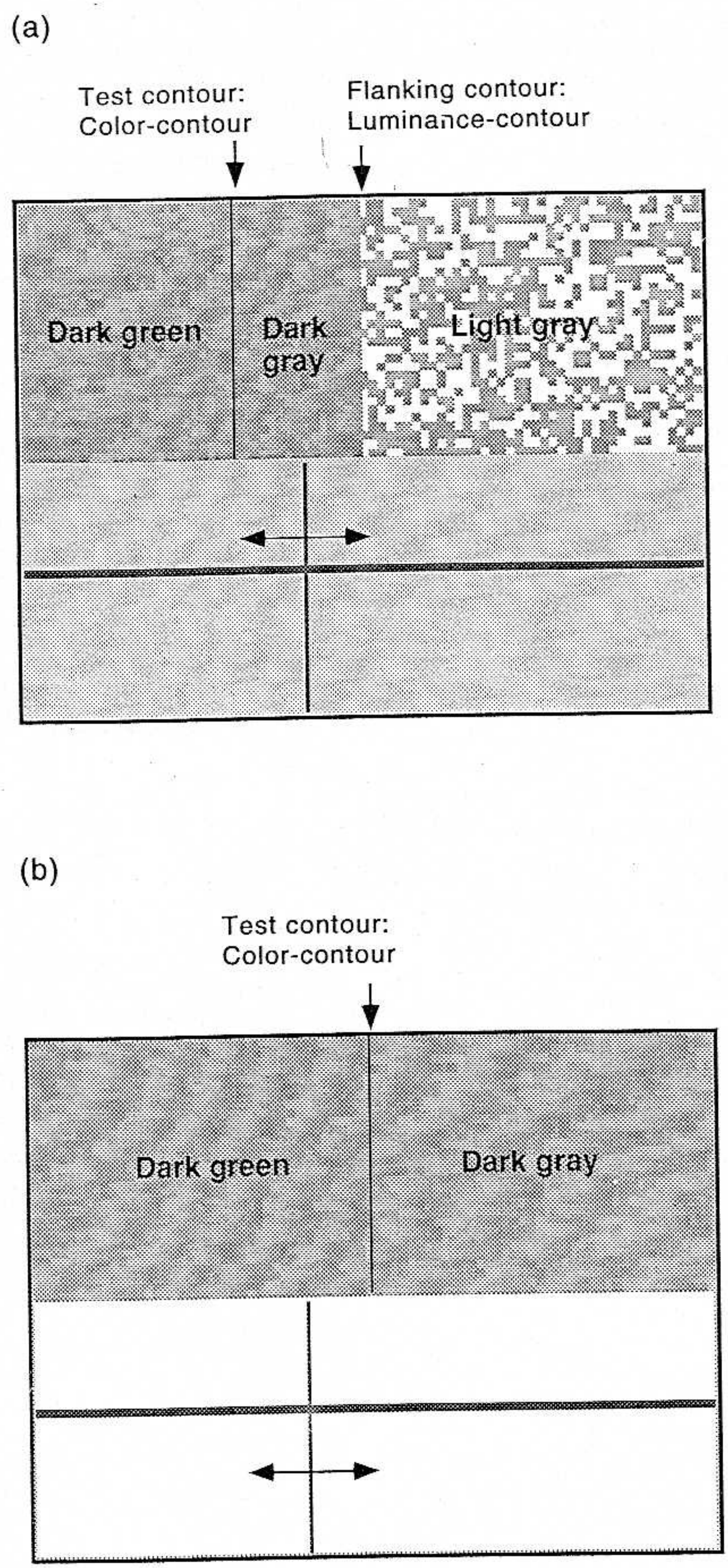

FIGURE 2. Illustration of the experimental display. In condition (a) the test contour was defined by color and the flanking contour was defined by luminance; in its corresponding baseline condition (b), only the test contour defined by color was presented. The random dot texture was dynamic in the real display. The comparison line was adjusted by the observers such that it appeared colinear with the test contour.

may have different appearances: the moving side would show some streaks along the direction of motion. Nevertheless, two factors argue against their importance. First, the streaks would on average be short and we found that texture cues are themselves the weakest cues for contours. We had to lower the contrast of all other cues to match the performance for texture (and our texture cue was very much like that which would be produced by time-averaging our motion cue). Second, the time averaging will necessarily lower the effective contrast of the textures making them even weaker. The comparison of the twinkling versus moving fields was our strategy for minimizing residual texture cues in the motion case while retaining a constant comparison field for all stimuli (twinkling, achromatic dots).

Moreover, using dynamic texture avoids featural cues that might be available in a static texture. The problem with static texture is that the observer might be able to inspect it and localize a particular dot or dot cluster and then localize the abutting contour of that dot. For example, when a "texture contrast" is introduced, the dot size changes and the observer might be able to localize the first tall dot. This would artificially increase the precision of localization. The dynamic texture makes this more difficult because no dot remains present long enough to localize as an individual feature.

In all conditions, a thin black vertical line (luminance, 0.0 cd/m^2; CIE coordinates, $x = 0.332$, $y = 0.333$*) was presented as a comparison line at the bottom half of the display. Its position was adjustable by moving the computer's mouse from side-to-side. A horizontal line at 18.6 min arc below the bottom of the random dot texture was presented as a fixation line. The intersection of the comparison line and the fixation line was the fixation point. Presenting the fixation point below the contour represents a precaution to ensure that precision of localization could be low even with luminance such that precision could be made about equal with other attributes. The background was white (luminance, 86.7 cd/m^2; CIE coordinates, $x = 0.332$, $y = 0.333$).

Procedure

All six pairs of attributes between luminance, color, texture, and motion were presented. For one pair (say, luminance and color), there were two conditions. In one condition, the test contour was defined by one attribute (e.g. color) and the flanking contour was defined by the other attribute (e.g. luminance). In the other condition, the attributes defining the test and the flanking contours were reversed. For each condition, the test contour was also presented without a flanking contour (baseline condition). For example, the condition—color test contour and luminance flanking contour—is illustrated in Fig. 2(a) and its corresponding baseline condition—a color test contour alone—is illustrated in Fig. 2(b): the left side of the display was unsaturated green equiluminant with the right side and the mean luminance of both sides was set at the luminance value needed to create the luminance flanking contour. In all conditions, the flanking contour was presented at 11 distances from the test contour (2.1, 4.3, 6.5, 10.8, 13.0, 15.1, 21.6, 28.1, 34.6, 41.0, and 47.5 min arc apart). Twenty-two adjustments were made at each relative distance, half when the test contour was to the left of the flanking contour and half when it was to the right. The order of presentation of the relative distances between the test and flanking contours and the order of presentation of the pairs of attributes were randomized for each observer.

In the control condition, both the test and flanking contours were defined by luminance (both contours had 40% contrast) and there was no random texture. The test contour was the darker one. The flanking contour was presented at 14 relative distances from the test contour (from 0.5 to 52.0 min arc). The test contour was presented without the flanking contour in one baseline condition.

The method of adjustment was used to measure the apparent position of the test contour. Observers were asked to adjust the middle of the comparison line presented in the bottom half of the screen to look colinear with the test contour. For example, in Fig. 2, observers adjusted the comparison line to look colinear with the color-contour. The minimum step size available with mouse movements was 0.24 min arc. The average adjusted position was taken as the perceived location.

The SE obtained for the 22 adjustments was taken as the measure of precision. The average precision across observers was 0.45, 0.42, 0.43, and 0.45 min arc, for luminance, color, texture, and motion respectively. (The corresponding SDs are 0.06, 0.04, 0.04, and 0.02.)

Results and Discussion

Method of analysis

To examine whether the location of the test contour is influenced by the position of the flanking contour, the mean location of adjustments taken at each condition was calculated and subtracted from the mean location obtained in its baseline condition. These differences are illustrated in graphs (see Fig. 3 as an example) where the distance between the flanking contour and the test contour is plotted on the x-axis, and the deviation from the mean location obtained in the baseline condition is plotted on the y-axis. On the y-axis, the zero point represents the mean location obtained in the baseline condition (straight line): a positive deviation shows that the test contour was perceived towards the flanking contour (attraction) and a negative deviation shows that the test contour was perceived away from the flanking contour (repulsion). A polynomial curve of the fifth order was fitted to the data. The maximum positive peak of the fitted curve represented the maximum attraction.

Control condition

Figure 3 illustrates the data obtained in the control condition in which both the test and the flanking contours were defined by luminance. The results show an interaction between the two luminance-contours

*0.0 cd/m^2 refers to the minimum luminance detectable by our photometer.

replicating the results found by Badcock and Westheimer (1985) and Rentschler *et al.* (1975). The test contour appeared towards the flanking contour when they were separated by 8.0 min arc or less. The maximum attraction happened when the test contour was at 1.8 min arc away from the flanking contour. At that relative distance, the test contour appeared shifted by 1.7 min arc towards the flanking contour. The maximum separation at which attraction was still found (8.0 min arc) is larger than the 3.0–4.0 min arc separation found by Badcock and Westheimer (1985). This difference may be due to the fact that we used a different experimental method. In one of Badcock and Westheimer's (1985) experiments, observers had to judge a jump of the test line as a flanking line was presented beside it and in another experiment, a vernier acuity task was performed using a two-alternative forced-choice (2AFC) task. Moreover, in both methods, the judgment of the location of the test line was done centrally. In the present experiment, the fixation point was presented 18.6 min arc below the test contour. Rentschler *et al.* (1975) presented a comparison line 30.0 min arc below the test bar and found attraction at separations of up to 10.0 min arc.

Even though there was a small tendency for repulsion in our data, no strong repulsion was found between the luminance contours. This result was not surprising since repulsion has not been consistently demonstrated among observers in the hyperacuity studies. As previously mentioned, Rentschler *et al.* (1975) showed a weak repulsion effect (a maximum shift of 0.3 min arc for a 30.0 min arc separation between the test and flanking line) for one observer while the other observer did not show any repulsion. Badcock and Westheimer (1985) consistently obtained repulsion in several observers using a 2AFC vernier task. However, differences in the magnitude of the effect between individuals do exist, and the effect was maximum with short presentation of the stimuli. Badcock and Westheimer (1985) also showed that, unlike attraction, repulsion is independent of the luminance contrast between the test and flanking lines. As a consequence of these findings, they wrote that repulsion "is not simply a byproduct of receptive fields with luminance weighting functions that contain an inhibitory surround" (p. 1266). It is possible that repulsion would be found for other observers or when using a 2AFC method with short presentation time in the present control condition.

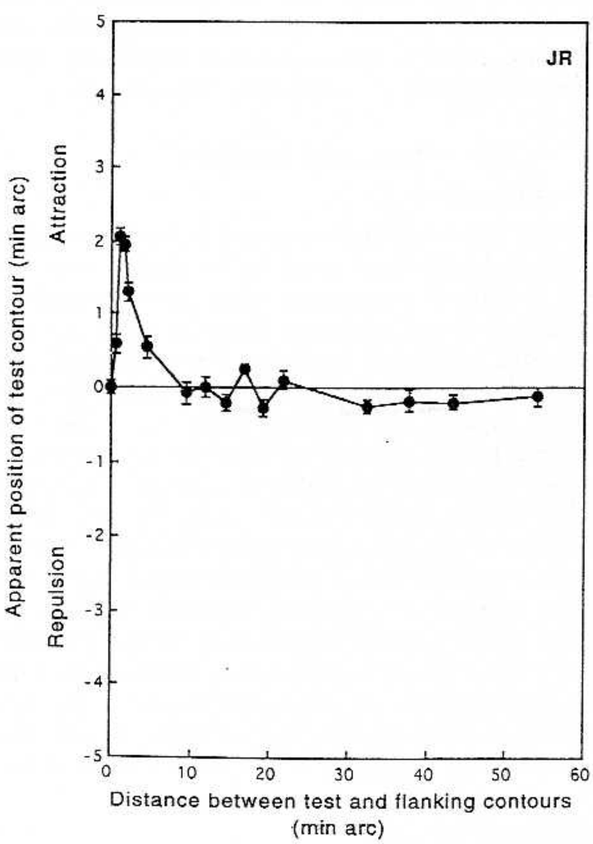

FIGURE 3. Results obtained in the control condition in which the test and flanking contours each had 40% luminance contrast.

Inter-attribute conditions

Interactions between the perceived position of adjacent contours defined by different attributes were found, showing that contour information from different attributes can interact at some site before a location decision is made. Moreover, a similar amount of interaction between different attributes was found showing that each attribute has a similar contribution to localization of contours at this common representation.

The patterns of attraction were similar for all observers. Therefore, the results were averaged across observers and a polynomial curve of the fifth order was fitted to these averaged results. The averaged results for all combinations of attributes are presented in Fig. 4: each graph has a format identical to that of Fig. 3. The maximum amount of attraction for each combination of attribute is given in min arc in the upper left corner of each graph. The results are similar across all combinations of test and flanking contours. Except when the test contour is defined by motion and the flanking contour is defined by texture, the results show consistent attraction. This interaction between different attributes shows that information from different attributes is combined at a common site.

In order to establish if luminance contributes more to localization of contours at that common location, the average amount of attraction produced by one attribute on all other attributes was calculated. On average, luminance attracted the other attributes by 0.6 min arc, color attracted them by 0.8 min arc, texture by 0.7 min arc and motion by 0.9 min arc. It is clear that luminance does not determine localization of contour since other attributes attracted luminance as much as the reverse or even a little more. Color attracted luminance by 1.0 min arc which is as much as the reverse (luminance attracted color by 1.0 min arc). Motion and texture attracted luminance by 1.0 and 1.4 min arc respectively, which is a little more than the reverse (luminance attracted motion and texture by 0.3 and 0.5 min arc respectively). From these results it can be

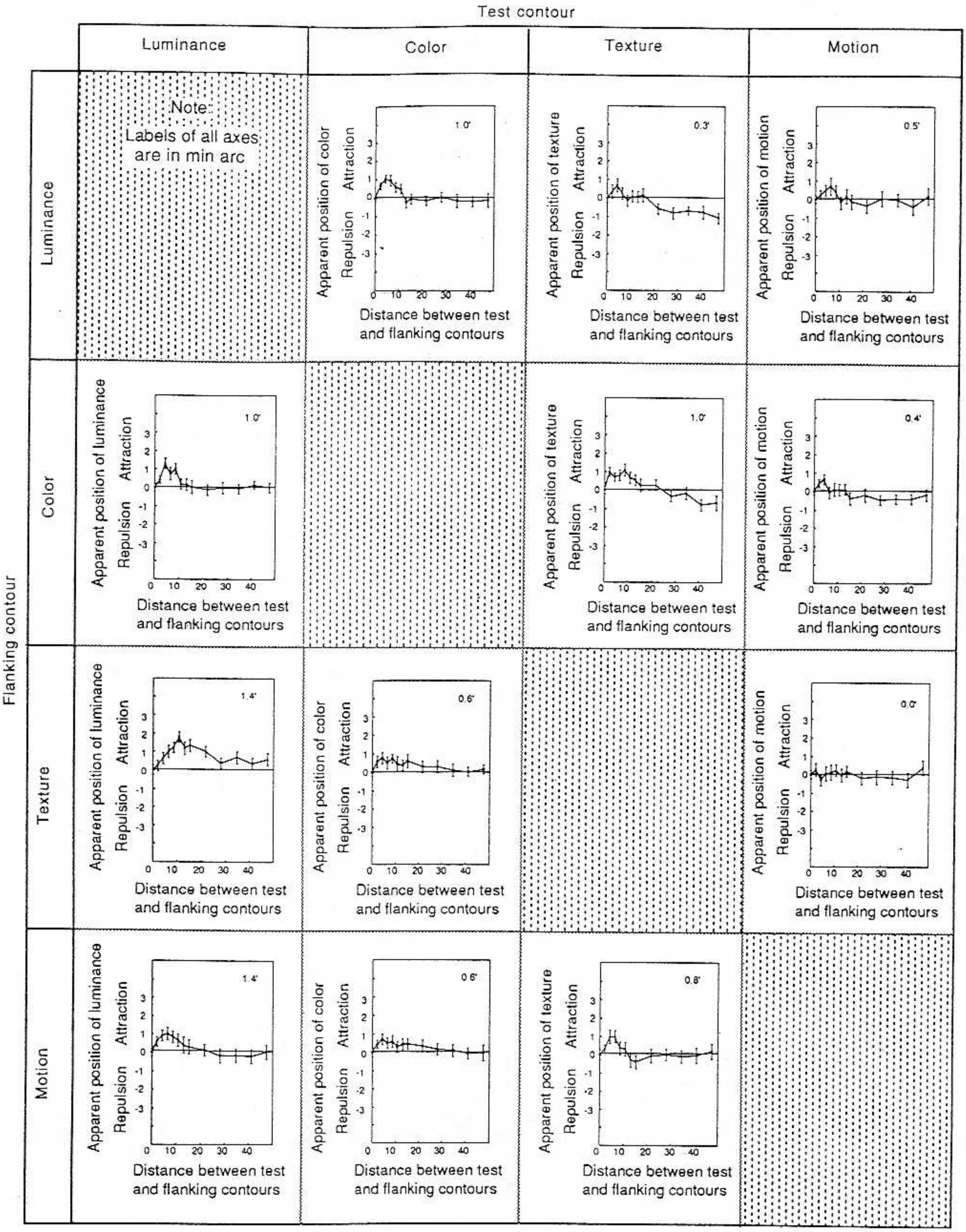

FIGURE 4. Results obtained in all inter-attribute conditions. The attribute that defined the test contour is presented on top and the attribute that defined the flanking contour is presented on the side. The format of each graph is identical to Fig. 3.

concluded that when attributes are presented at a contrast that produced about the same precision of localization, luminance does not have undue influence over other attributes.

In general, the interaction of one attribute with another was reciprocal. However, some asymmetries in the pattern of interactions were found for pairs of attributes including texture. A texture-contour did not attract a color-contour as much as the reverse. In addition, a texture-contour did not attract a motion-contour at all, whereas there was attraction in the reverse case. It would be premature to draw a strong conclusion about the specific contribution of texture in localization since texture could be defined in many different ways and our results may only apply to the particular texture we chose.

Only a slight tendency towards repulsion was found. Repulsion effects were much weaker and more variable across combinations of attributes and across observers than attraction effects. This weak effect and variability in repulsion are consistent with the results found in the literature on acuity (Badcock & Westheimer, 1985; Rentschler *et al.*, 1975). Despite the weak repulsion and the weakness of texture, the consistent influence of one attribute on another one shows that localization of contours originates from a site at or beyond where activity profiles from different attributes combine.

In addition, the results show that the relative distance at which contours defined by different attributes interact is about the same no matter what pair of attributes was studied (most test contours were attracted by a flanking contour when they were separated by 10.5 min arc or less). This may be a consequence of selecting "equal precision" contrasts for each attribute.

Our conclusion differs from that proposed by Banton and Levi (1993): they argue that there are independent localization mechanisms for motion-defined and luminance-defined targets. Our results do not rule out this possibility, they rather suggest that luminance and motion information further combine at a common site. Moreover, the methodology used by Banton and Levi (1993) differs greatly from our methodology. First, they measured vernier threshold (a measure of precision of localization) whereas we measured apparent location. Badcock and Westheimer (1985) suggested that the coding of precision and the coding of location may require different processing mechanisms. The difference in our results may represent a support for Badcock and Westheimer's suggestion; directly comparing the measure of precision and location between the luminance and motion mechanisms would be essential to clarify the issue. Second, the random dot density was identical for our luminance- and motion-contours, whereas in Banton and Levi's study, the luminance- and motion-bars were not defined, with dots having the same density. Third, the spatial location judgment was done at 18.6 min arc in our experiment but it was done foveally in Banton and Levi's experiment.

EXPERIMENT 2

Experiment 2 examined whether precision to localize a contour defined by more than one attribute is improved and whether this improvement is a consequence of the combination of information at a common neural site. Statistically, if separate measurements about the same contour are available each having independent noise, combining these measures at a common site improves precision. Imagine again that a contour is represented by a simply-peaked distribution of neural activity on a cortical surface (see Fig. 5) and that location decision depends on the position of the peak of the profile of activity (the mean or mode of the distribution). When the two profiles of activity that signal that the same position are summed, the SE (SD divided by square root of the number of measurements) of the mean of the summed profile of activity will be smaller. This SEM corresponds to the precision of localization, thus precision should improve as the number of attributes superimposed increases. This is true assuming that the profiles of activity associated with each attribute are summed and that their noise is independent. The

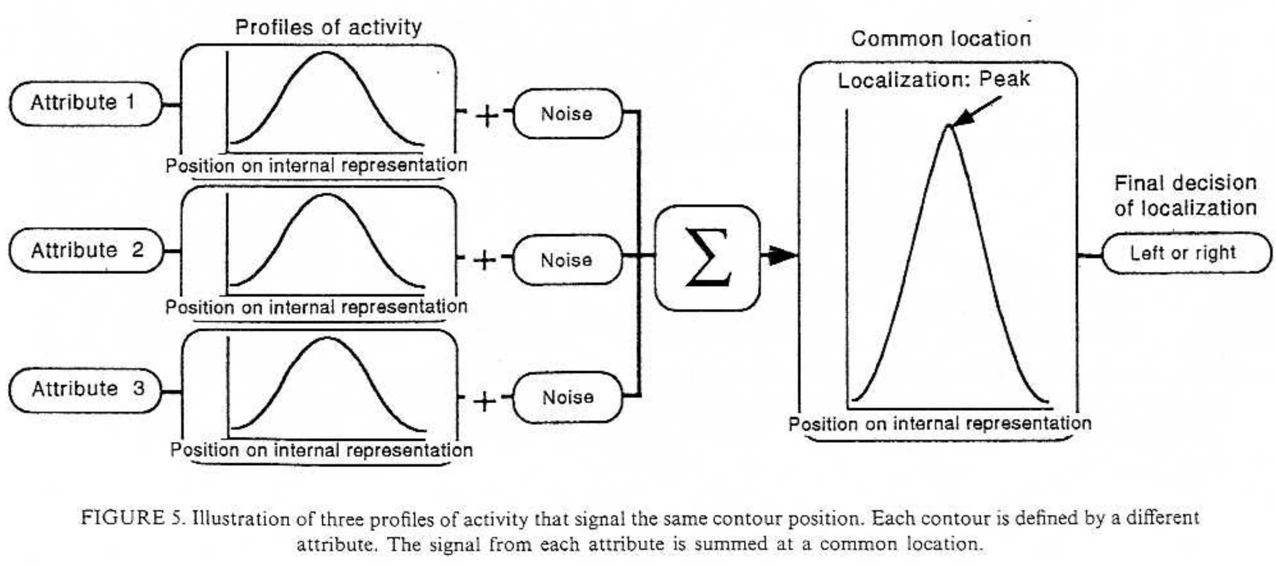

FIGURE 5. Illustration of three profiles of activity that signal the same contour position. Each contour is defined by a different attribute. The signal from each attribute is summed at a common location.

signal-to-noise ratio is given by $1/\sqrt{n}$ for n attributes if the signal from each attribute has equal noise variance and is weighted equally. For example, when two attributes define the contour, the precision should improve by 30% $(1 - 1/\sqrt{2})$ and when three attributes define the contour, it should improve by 42%, if each attribute provides equal precision and is equally weighted.

Of course, if several attributes are combined but some offer a better precision than others, the combined precision would improve only if the visual system gives advantages to the attributes that offer the highest precision in the image and de-emphasizes others. For example, imagine a contour whose position in luminance is not sharply defined but whose position in color is sharply defined. The color information can effectively be used for localization whereas the luminance information hinders the localization. If the visual system simply adds the two representations together, it ends up with a combined representation that is worse than the representation in color only. For maximum improvement, the summation should give color a larger weight than luminance.

In Expt 2, precision to localize contours defined either by luminance, color, texture, or a combination of two or three of them was measured using a 2AFC procedure. In the single-attribute conditions, precision was measured when only one attribute defined the contour: a vertical test contour was created by varying the luminance, the color or the texture of one side of the display. In the combined-attributes conditions, attributes were superimposed in pairs or trio. To avoid the possibility that the visual system might weight attributes according to their quality in the image, each attribute was presented at a contrast that gave about equal precision of localization.

A control condition was run in order to determine the best precision possible using our display: precision to localize a 100%-contrast luminance-contour was measured.

Method

Subjects

The same three observers as in Expt 1 participated. Only observer JR was tested in the control condition.

Apparatus

A Macintosh II cx was used. The experimental display (88.0 min arc height × 116.0 min arc width) was presented on an Apple color monitor. Observers were seated with their head and chin supported by a rest.

Except in the control condition, the top half of the display was filled with dark and light dynamic random dot texture. This texture was identical to the dynamic texture used in Expt 1 except that the square dots were 1.5 min arc on a side and the texture was replaced every 45.0 msec.

As in Expt 1, the properties of the texture on the left side of the display were varied such that a vertical contour was created, however only one contour was presented at one time. In the single-attribute conditions, the test contour was defined by luminance, color or texture. In the combined-attributes conditions, the test contour was defined by two (e.g. color and luminance: darker green on the left and gray on the right), or three (luminance, color and texture) attributes. All possible pairs of attributes were studied [See Fig. 2(b) for an example of a color test contour.] As in Expt 1, the dynamic texture was used to ensure that precision of localization could be low for luminance alone; thus it could be about equal for each attribute and room was left for improvement when attributes were superimposed.

The "luminance contrast", "color contrast' ' and the "texture contrast" were defined identically to those used in Expt 1. The exact values used for each observer and each condition are given in the following paragraphs.

Luminance. A decrease in mean luminance of 15.0% was needed for observer JR to detect the contour. Decreases of 20.0% and 17.5% were needed for observers TWB and ZH respectively.

Color. The green saturation was at about 16.0% for all observers.

Texture. For all observers, the "texture contrast" creating a texture contour was a 400% increase in the height of the dots on the left side (height of 5.9 min arc and width of 1.5 min arc) compared to those on the left side while all other properties of the two textures were equal.

In the combined-attributes conditions, the two or three attributes were superimposed using the contrast values presented above.

In the control condition, the top half of the screen did not have random dot texture, and the vertical test contour was defined by a luminance discontinuity of 100% contrast (luminance, 0.0 cd/m^2 on the left, 40.0 cd/m^2 on the right; CIE coordinates $x = 0.332$, $y = 0.333$).

The bottom half of the display was identical to the one used in Expt 1.

Procedure

A vernier acuity task was used. Precision of localization was determined using a 2AFC procedure. The test contour was randomly presented to the left of the middle of the screen for half of the trials and to the right for the other half. The comparison line was vertically aligned with the test contour, or displaced laterally by a maximum of 4.0 min arc to the right or left of the test contour. For observer ZH, the comparison line was presented at 2.6, 1.8, 1.5, 1.1, 0.7, and 0.4 min arc, to the left and right of the test contour and at 0.0 min arc directly colinear with the test contour. For observer JR, the comparison line was presented at 1.8, 1.5, 1.1, 0.7, and 0.4 min arc to the left of the test contour, and at 3.3, 2.6, 1.8, 1.5, 1.1, 0.7, and 0.4 min arc to the right of and colinear to the test contour. For observer TWB, the comparison line was presented at 3.3, 2.6, 1.8, 1.5, 1.1, 0.7, and 0.4 min arc to the left of the test contour and at 1.8, 1.5, 1.1, 0.7 and 0.4 min arc to the right of and colinear to the test contour. These values bracket the

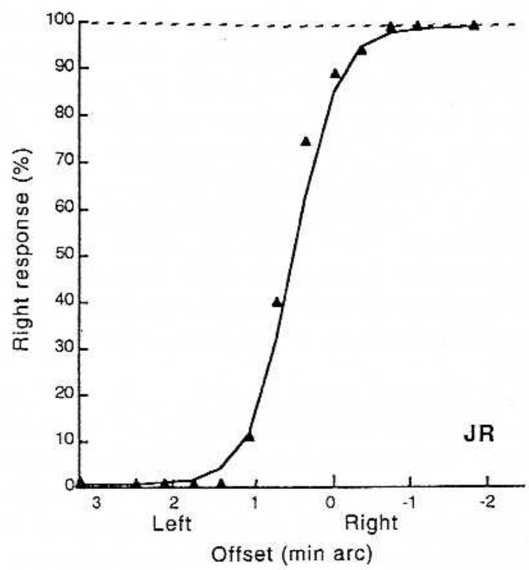

FIGURE 6. Psychometric function obtained with the 100%-contrast luminance contour used in the control condition.

mean localization setting for each observer as determined in a pilot run. For all observers, 20 measurements were taken at each position.

Observers fixated at the intersection of the comparison and fixation lines, and reported whether the position of the comparison line was to the left or to the right of the vertical test contour. The stimulus stayed on until the observer responded and no feedback was given.

Results and Discussion

To examine the precision of localization, a psychometric function was obtained for each condition and observer. The distance between the comparison line and the test contour to the left or to the right in min arc is plotted on the x-axis, and the percentage of trials in which the observers reported seeing the comparison line to the right of the test contour is plotted on the y-axis. A modification of the hyperbolic arctan function was fitted to the data using a least squares criterion.* The precision is given by the just noticeable difference (JND: half the difference between the position on the fitted function at which the test appeared to the right 25% of the time and the position at which it did so 75% of the time). The obtained precisions of localization when two or three attributes define the contour was correlated with precisions that should be obtained if there is summation of neural activities at a common site.

Control condition

Figure 6 shows the psychometric function obtained in the control condition (100%-luminance contrast) for observer JR. The results show a JND of 31.0 sec arc. This performance is lower than the usual hyperacuity of 2.0–5.0 sec arc when the contour is presented foveally at a high contrast luminance. Precision is undoubtedly lower because the judgment of the relative position between the test contour and the comparison line was performed at 18.6 min arc from the fovea.

Figure 7 illustrates the JNDs that were derived from the psychometric functions obtained by each observer for each single-attribute condition and for each combined-attributes condition.

Single-attribute conditions

For all observers, precisions obtained when only one attribute was presented were substantially worse than the best precision of 31.0 sec arc found in the 100%-luminance control condition. On average across observers, the precision of localization was 58.0 sec arc for luminance, 60.0 sec arc for color, and 78.0 sec arc for texture. Therefore, some margin for improvement was available when the attributes were presented together.

On average, the precision for localizing a texture-contour was lower than that for localizing luminance- and color-contours. It was nevertheless the best performance obtainable with the range of textures used in our stimuli.

Combined-attributes conditions

The average JNDs across observers are 49.0 sec arc for the combined texture–color contour, 41.0 sec arc for the color-luminance contour and 59.0 sec arc for the texture–luminance contour. The average JND across observers is 38.0 sec arc when all three attributes were superimposed, approaching the 31.0 sec arc found with the 100%-luminance contrast. These results show that precision improves as the number of attributes defining the contours increases.

The precision that predicted if there is summation of the neural activity related to each contour at a common site was calculated using the precision (JND) obtained by each observer with each attribute alone. The predictions are made assuming a neural summation of the profile of activity associated with each attribute in the image (as illustrated in Fig. 5). In this summation, it is assumed that the SE of the profile of activity corresponds to the precision of localization—the JNDs in our experiment. It is also assumed that the noise associated with each attribute is independent; therefore the

*A modification of hyperbolic arctan function was fitted to the data. The function was adapted to suit the axes of the graphs. Normally, minimum and maximum values of the y-axis of a hyperbolic arctan function are -1 and 1 respectively. However, since the percent of "right" responses is plotted along the y-axis of the graph, the minimum and maximum values must be 0 and 100 respectively. In order to have these limits on the y-axis, the following function was used to perform the least square fitting:

$$y = 50(\tan h(mx + b) + 1).$$

The tan h of $mx + b$ was used because the slope and ordinate of each fitted function need to vary according to the distribution of a different set of data. The data points of 0 and 100 were replaced by 1 and 99 respectively to perform the curve fitting because $y = 0$ and $y = 100$ are asymptotes of the tan h function, such that the function cannot reach these limits. In order to find the slope and ordinate of the distribution of each set of data, a tan h $(y/50 - 1)$ was applied on the percent of "right" responses obtained (y values) such that the y values were a linear function of the x values and a least square of a tan h $(y/50 - 1)$ and x values was obtained.

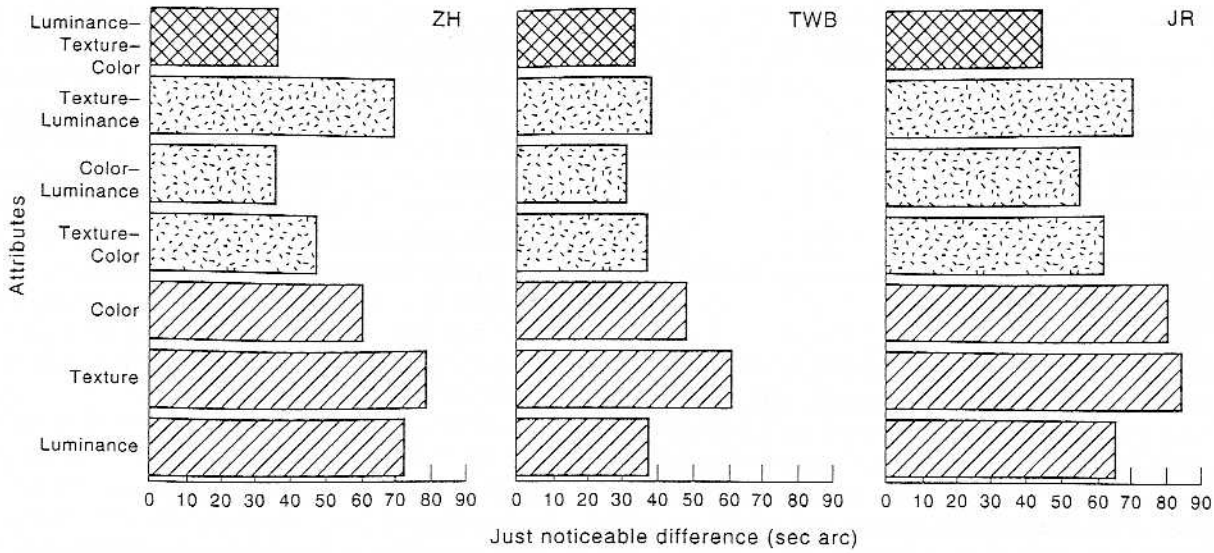

FIGURE 7. JNDs in sec arc for all observers and all combinations of attributes.

signal-to-noise ratio is given by $1/\sqrt{n}$. In addition, each attribute was given an identical weight even though texture alone afforded a precision lower than luminance or color. By assuming equal weight, the prediction of improvement will be less than expected if the visual system maximizes the combination of attributes. Maximizing the combination requires giving a smaller weight to the attribute that affords the worst precision (in this case, texture).

For example, when the three attributes—luminance, color and texture—were superimposed, formula (1) was used to calculate the predicted precision for each observer separately:

$$\mathrm{JND}_{(3)} = \frac{\sqrt{\dfrac{\sigma^2_{\mathrm{lum}} + \sigma^2_{\mathrm{col}} + \sigma^2_{\mathrm{text}}}{n}}}{\sqrt{n}} \qquad (1)$$

where σ is the JND obtained when the contour was defined by the attribute named in subscript (i.e. luminance, color, or texture), and n is the number of attributes defining the contour (i.e. three).

The obtained precisions are plotted against the predicted precisions for each observer and each combination of attributes in Fig. 8. Despite the fact that these predicted precisions are not adjusted to optimize the combinations, the results are positively correlated with the predictions ($r = 0.88$, $P < 0.001$).

In addition, using the obtained precision for each attribute alone, the percentage improvements predicted for two and three attributes were calculated. These predicted improvements are compared to the obtained ones in Table 1.

The similarity between the obtained and predicted results supports the model described above where the final decision of localization happens at a neural site following the summation of information from all attributes. In agreement with this conclusion, Frome, Buck and Boynton (1981) showed that information from color and luminance combined to improve visibility of a border.

Moreover, such a strong positive correlation between the obtained and predicted precisions shows that the visual system did not give an advantage to the attribute that afforded the best precision, and neither did it de-emphasize the attribute that afforded the worst precision. For example, for observer JR, luminance offered the best precision and texture offered the worst; when these two attributes were combined the resulting precision was worse than the one with luminance alone. Localization would have been more precise if texture was simply ignored by the visual system, but this did not happen. These results suggest that when combined, the information associated with a given attribute is not weighted according to the precision that this attribute can afford when presented alone.

Cavanagh, Tyler and Favreau (1984) also found results suggesting that the visual system does not give a

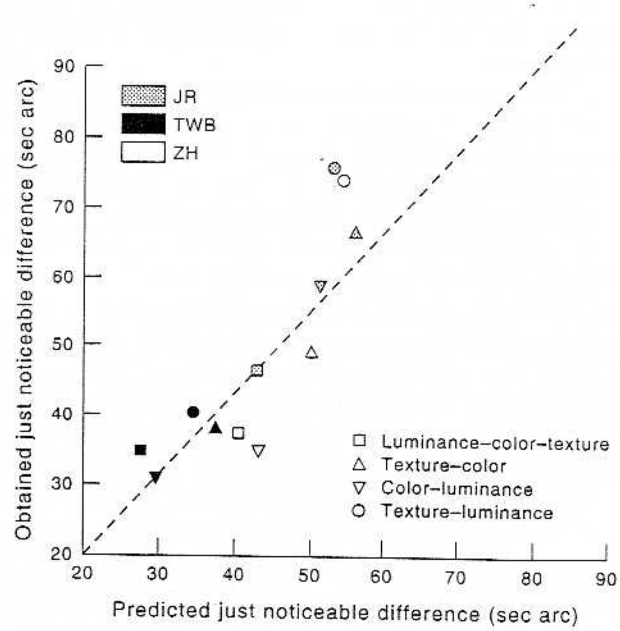

FIGURE 8. Precisions (JNDs) predicted if the attributes are combined at a common location or each observer and each combination of attributes as a function of obtained precision.

TABLE 1. Average percentage improvements of precision across observers for the three pairs and the trio of attributes

Attributes	Improvement (%)	
	Obtained	Predicted
Luminance–color	30.5	29.3
Color–texture	27.8	27.5
Texture–luminance	12.2	26.9
Luminance–color–texture	40.8	40.7

smaller weight to an attribute which hinders performance. They measured the perceived velocity of a luminance grating by adding chrominance modulation to it and found that its perceived velocity was slowed down. They assumed that perceived velocity is derived from a sum of the separate color and luminance analyses. In this summation, even if color analysis signals a much lower velocity than luminance analysis, it was not given a smaller weight. Our results are similar to their results; attributes offering worse precision of localization were nevertheless included in the decision. It could be argued that the visual system could not know the precision of an attribute in a given condition although it might accumulate such knowledge over time. Such long-term world knowledge did not seem evident here either as luminance, which should certainly accumulate the best record, did not demonstrate any inherent dominance. These results imply at the least that contour localization is not necessarily determined by luminance as claimed by Gregory and Heard (1979), Grossberg and Mingolla (1985), Livingstone and Hubel (1984), and Yeh *et al.* (1992).

In short, the results show that the obtained precisions correlated positively with precisions predicted from a model of neural summation in which information from different attributes is summed at a common site and where each has an equal contribution (given that their contrasts have been set to produce similar precision in isolation).

One could argue that even if information from different attributes never combines at a common site, precision could improve simply due to the probability of improving the decision by accumulating separate decisions, each based on a different attribute. However, the obtained precisions greatly differ from those that are predicted from probability summation.

In order to understand how precision would improve from probability summation in a 2AFC procedure, we must first model how a detection task is improved by probability summation. In a detection task, observers must decide whether they see a stimulus or not. It is clear that, in this task, the probability of reporting the stimulus will be increased when it is defined by many types of signals. Indeed, the chance that any one of the signals exceeds threshold is always greater than the chance that a given signal exceeds threshold, thus causing improvement of detection (e.g. see Green & Swets, 1974; Pelli, 1985, for a review of probability summation for detection tasks).

In a localization task, however, observers must do more than detecting the contour; they must somehow report its position. In a 2AFC procedure—where observers reported the position of the test contour by saying whether a comparison line is to its left or to its right—the probability of the separate outcomes of independent decisions, one for each attribute, would make precision of localization better as attributes are superimposed, but in a manner less intuitive than in a detection task.

Unlike in a detection task, where a single yes vote from any one attribute is sufficient to indicate the presence of the signal, a majority vote is required in a discrimination task. Imagine that a contour is defined by two attributes. What response is given if one attribute votes for "left" and the other for "right"? Assuming that ties are resolved randomly, no improvement can result from combining two independent decisions. On the other hand, when three attributes are superimposed, there are no ties and a clear and predictable 30% improvement results. For example, two "lefts" and one "right" are sufficient to decide "left" and so less signal is required than that which produces three independent "lefts" in isolation (the baseline). This pattern of prediction is distinctly different from the observed results and the "majority vote" probability summation model is rejected. If the information being combined in the probability summation were some continuous probability estimate (e.g. 57% chance of left) rather than a final decision (left or right), then a suitable probability summation model could predict the results of Fig. 8. However, our first experiment showed that the continuous variable being summed across attributes was, if anything, the profile of activity from that attribute in response to the contour. Our model of statistical noise reduction above is also based on summing of activity profiles (now aligned) so overall we feel that this is the most parsimonious explanation for both experiments.

CONCLUSIONS

The results of Expts 1 and 2 suggest that information from different attributes is united at a common site to provide localization of contours. Experiment 1 showed that the position of one contour defined by one attribute influenced the position of a contour defined by another attribute. Spatial interaction between contours defined by different attributes show that responses to the different attributes must reach a common location prior to the localization decision. The patterns of interaction further suggest that no attribute predominates in determining localization and that, in fact, all attributes have similar contributions. These conclusions were supported by the results of Expt 2; precision to localize contours changed as the number of attributes defining the contour increased. The amount of changes found are consistent with summing activity profiles from different attributes at a common location.

Where would this common site be in the visual cortex? Results from single cell recordings suggest that it can be as early as the visual area V4. Ferrera, Nealey and

Maunsell (1991), Maunsell, Nealey and Ferrera (1992), and Logothetis and Charles (1990) found cells selective for more than one attribute in area V4. In this area Logothetis (personal communication) found cells selective for the orientation of contours defined by several different attributes. Logothetis and Charles (1990) concluded: "The results suggest an integration of visual cues, at least at the level of area V4, for the extraction of shape information."

Our results show that combining information from different attributes may be a strategy used by the visual system to enhance precision of localization. When attributes presented at a contrast which produced approximately equal precision of localization are superimposed, precision to localize contour was improved. These results are in agreement with Morgan's (1986) results showing that our precision to localize a contour defined by luminance and disparity is better than our precision to localize a contour defined by either one alone. Combining information from different attributes is a strategy that also enhances visual analyses other than contour localization. Goodale, Humphrey, Milner, Jakobson, Servos and Carey (1991) showed that object recognition is facilitated when the objects are defined by a greater number of attributes. Frome *et al.* (1981) showed facilitation for the detection of borders, Bülthoff and Mallot (1988) showed increased sensitivity for depth perception, and Treisman and Sato (1990) showed that visual search is faster as the number of attributes defining the searched stimulus increases.

Our results do not support the argument that luminance affords a localization advantage over other attributes (Gregory & Heard, 1979; Grossberg & Mingolla, 1985; Livingstone & Hubel, 1984; Yeh *et al.*, 1992). After removing the advantage of its high spatial resolution at high contrast, we found that luminance does not play a privileged role in localization of contours. Conversely, luminance may play a privileged role in localization of contours in natural scenes because it typically has a very high contrast compared to that attainable for other attributes. That high contrast produces more accurate localization than can other attributes. For example, Yeh *et al.* (1992) studied spatial localization using a figural aftereffect paradigm in which luminance and color stimuli were presented. However, they presented each attribute at an equivalent multiple of the contrast needed for detection. They showed a strong figural aftereffect between an inducing stimulus defined by luminance and a test stimulus defined by color; however the aftereffect was reduced when the inducing stimulus was defined by color and the test stimulus defined by luminance. They concluded that "luminance may have more weight than chrominance on the 'gluing' of spatial position" (p. 704). In light of our results, we believe that given the contrasts used in their display, the precision of localization would have been substantially better with luminance than with color. If the color- and luminance-defined stimuli had been presented at contrasts that produced equal precision of localization, as in our experiments, the strength of their aftereffect may have been more symmetrical in both experimental conditions.

We believe that all attributes may play an essential role in localization of contours. For example, pooling information from different attributes may be advantageous to understanding scenes with shadows. Discontinuities in luminance created by shadows are not reliably linked to the contours of objects, whereas discontinuities in other attributes (e.g. color, motion, and texture) are much more reliably linked to object contours. This advantage of multiple analyses has been exploited in computer vision by Crissman (1990). She designed a system that used image hue and luminance data to navigate an automated land vehicle on natural roads. Because her system uses color contrast in addition to luminance contrast in the image, it has the advantage of being able to pick out the road contours even in heavily shadowed scenes. This advantage of multiple analyses should be further studied in psychophysical experiments.

REFERENCES

Albright, T. D. & Desimone, R. (1987). Local precision of visuotopic organization in the middle temporal area (MT) of the macaque. *Experimental Brain Research*, *65*, 582–592.

Anderson, S. J. & Burr, D. C. (1987). Receptive field size of motion detector units. *Vision Research*, *27*, 621–635.

Badcock, D. R. & Westheimer, G. (1985). Spatial location and hyperacuity: the centre/surround localization contribution function has two substrates. *Vision Research*, *25*, 1259–1267.

Banton, T. & Levi, D. M. (1993). Spatial localization of motion-defined and luminance-defined contours. *Vision Research*, *33*, 2225–2237.

Botez, M. I. (1975). Two visual systems in clinical neurology: Readaptative role of the primitive system in visual agnosic patients. *European Neurology*, *13*, 101–122.

Bülthoff, H. H. & Mallot, H. A. (1988). Interaction of depth modules: Stereo and shading. *Journal of the Optical Society of America*, *5*, 1749–1758.

Cavanagh, P., Tyler, C. W. & Favreau, O. E. (1984). Perceived velocity of moving chromatic gratings. *Journal of the Optical Society of America*, *1*, 893–899.

Crissman, G. (1990). Color vision for the detection of unstructured roads and intersections. Ph.D. thesis, Electrical and Computer Engineering Department, Carnegie-Mellon University, Pittsburg, Pa.

Damasio, A., Yamada, T., Damasio, H., Corbett, J. & McKee, J. (1980). Central achromatopsia: Behavioral, anatomical and physiologic aspects. *Neurology*, *30*, 1064–1071.

Day, R. H. (1962). Excitatory and inhibitory processes as the basis of contour shift and negative after-effect. *Psychologia*, *5*, 185–193.

van Essen, D. C. & Maunsell, J. H. R. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in Neuroscience*, *6*, 370–375.

Ferrera, V. P., Nealey, T. A. & Maunsell, J. H. R. (1991). Magnocellular and parvocellular inputs to macaque area V4. *Investigative Ophthalmology and Visual Science*, *32*, 1117.

Frome, F. S., Buck, S. L. & Boynton, R. M. (1981). Visibility of borders: Separate and combined effects of color differences, luminance contrast, and luminance level. *Journal of the Optical Society of America*, *71*, 145–150.

Ganz, L. & Day, R. H. (1965). An analysis of the satiation-fatigue mechanism of figural after-effects. *American Journal of Psychology*, *128*, 345–361.

Gibson, J. J. (1933). Adaptation, after-effect and contrast in the perception of curved lines. *Journal of Experimental Psychology*, *16*, 1–31.

Goodale, M. A., Humphrey, G. K., Milner, D. A., Jakobson, L. S., Servos, P. & Carey, D. P. (1991). Object versus picture identification in a patient with visual form agnosia. *Investigative Ophthalmology and Visual Science*, *32*, 1181.

Green, D. M. & Swets, J. A. (1974). *Signal dectection theory and psychophysics*. Huntington, N.Y.: Krieger.

Gregory, R. L. (1977). Vision with isoluminant colour contrast: 1. A projection technique and observations. *Perception*, *6*, 113–119.

Gregory, R. L. & Heard, P. (1979). Border locking and the Café Wall illusion. *Perception*, *8*, 365–380.

Grossberg, S. & Mingolla, E. (1985). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, *92*, 137–211.

Hines, M. (1976). Line spread function variation near the fovea. *Vision Research*, *16*, 567–572.

Hock, H. S. & Eastman, K. E. (1995). Context effects on perceived position: Sustained and transient temporal influences on spatial interactions. *Vision Research*, *35*, 635–646.

Landy, M. S. (1993). Combining multiple cues for texture edge localization. In Rogowitz, B. E. & Allebach, J. P. (Eds), *Human vision, visual processing and digital display IV. Proceedings of the SPIE*, *1913*, 506–517.

Livingstone, M. S. & Hubel, D. H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience*, *4*, 309–356.

Livingstone M. & Hubel, D. H. (1988) Segregation of form, color, movement and depth: Anatomy, physiology and perception. *Science*, *240*, 740–749.

Logothetis, N, K. & Charles E. R. (1990). V4 responses to gratings defined by random dot motion. *Investigative Ophthalmology and Visual Science*, *31*, 90.

Maunsell, J. H. R. & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, *10*, 363–401.

Maunsell, J. H. R., Nealey, T. A. & Ferrera, V. P. (1992). Magnocellular and parvocellular contribution to neuronal responses in monkey visual cortex. *Investigative Ophthalmology and Visual Science*, *33*, 901.

Mollon, J. D., Newcombe, F. Polden, P. G. & Ratcliff, G. (1980). On the presence of three cone mechanisms in a case of total achromatopsia. In Verriest, B. (Ed.), *Colour vision deficiencies*, *V* (Chap. 3, pp. 130–135). Bristol: Hilger.

Morgan, M. J. (1986). Positional acuity without monocular cues. *Perception*, *15*, 157–162.

Pearlman, A. L., Birch, J. & Meadows, J. C. (1979). Cerebral color blindness: An acquired defect in hue discrimination. *Annals of Neurology*, *5*, 253–261.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A*, *2*, 1508–1532.

Regan, D., Giaschi, D., Sharpe, J. A. & Hong, X. H. (1992). Visual processing of motion-defined form: Selective failure in patients with parietotemporal lesions. *Journal of Neuroscience*, *12*, 2198–2210.

Rentschler, I., Hilz, R. & Grimm, W. (1975). Processing of positional information in the human visual system. *Nature*, *253*, 444–445.

Rovamo, J., Hyvärinen, L. & Hari, R. (1982). Human vision without luminance contrast system: Selective recovery of the red–green colour-contrast system from acquired blindness. *Documenta Ophthalmologica Proceedings Series*, *33*, 457–466.

Schiller, P. H. & Colby, C. L. (1983). The responses of single cells in the lateral geniculate nucleus of the rhesus monkey to color and luminance contrast. *Vision Research*, *23*, 1631–1641.

Treisman, A. & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 459–478.

Tyler, C. W. & Nakayama, K. (1984). Size interactions in the perception of orientation. In Spillman, L. & Wooten, B. R. (Eds), *Sensory experience, adaptation and perception*, (pp. 529–547).

Yeh, S.-L., Chen, I.-P., De Valois, K. K. & De Valois, R. L. (1992). Figural aftereffects with isoluminant gaussian blobs. *Investigative Ophthalmology and Visual Science*, *33*, 704.

Zihl, J., von Cramon, D. & Mai, N. (1983). Selective disturbance of movement vision after bilateral brain damage. *Brain*, *106*, 313–340.

Acknowledgements—This research was supported by NSERC grant 015593 and a Glendon College grant to JR and Air Force Office of Scientific Research 91-0169 to PC. The authors wish to thank the two reviewers for their valuable suggestions, Raynald Comtois for his technical support and James Intriligator, Isabelle Boutet, Hiroshi Ono and his lab members for their editing comments.