Texture segregation by chromatic and achromatic visual pathways: an analogy with motion processing

Andrei Gorea

Laboratoire de Psychologie Expérimentale, René Descartes University, Centre National de la Recherche Scientifique, 28 rue Serpente, 75006 Paris, France

Thomas V. Papathomas

Laboratory of Vision Research and Department of Biomedical Engineering, Rutgers, the State University of New Jersey, Piscataway, New Jersey 08903

Received August 15, 1989; accepted May 29, 1990

We present results to show that texture segregation can be obtained through the so-called coherent spatial grouping of local shape (orientation) and of local color under both nonequiluminant and equiluminant conditions. Color grouping entails texture segregation independent of orientation grouping, while the reverse is not true under equiluminant conditions. The experiments permit the isolation of chromatic- and luminance-oriented mechanisms, as well as of chromatic nonoriented mechanisms, all of which contribute to texture discrimination. As a general rule, the present results (including the asymmetry between color and orientation grouping) are similar to those obtained by us in a series of motion-perception experiments. This similarity suggests that the perceptual rules governing spatial grouping are analogous (if not identical) to those governing spatiotemporal grouping. As in the case of directional discrimination, texture-discrimination performances may be accounted for by the activation of higher-order units receiving inputs from subunits, all of which display similar tuning properties within a multidimensional space.

INTRODUCTION

The display in the x–t plane (where time t is shown as y, the vertical axis) of a one-dimensional moving stimulus is a convenient way to represent motion trajectories.[1] There is a wide variety of moving stimuli whose spatial representation in the x–t plane appears as a characteristic textured surface (e.g., Figs. 16 and 17 in Ref. 1 and Figs. 1 and 4 in Ref. 2). The general question asked in this paper is whether the rules governing motion perception do have a counterpart in the domain of texture segregation.

The above question is not gratuitous. In a series of recent studies concerned with the perception of apparent motion elicited by the correspondence matching of multiple image attributes,[3–7] we used a class of stimuli that permitted any attribute of the visual targets to be arranged spatiotemporally such that they may elicit motion perception in a selected direction independently of the remaining attributes. We were thus able to produce apparent-motion sequences in which motion was elicited by coherently matching the color of the targets, while their orientation was either fixed (i.e., shared by all elements) or mismatched across frames. In a dual manner, we also produced sequences in which the orientation of the elements was matched across frames in order to attempt to produce so-called orientation-based coherent motion; simultaneously, the color of the elements was either fixed or mismatched across frames. We found that color matching produced coherent motion perception whether orientation was fixed or mismatched with both dark and equiluminant (to the elements) backgrounds. By contrast, when the background was equiluminant, orientation matching elicited weak motion perception for unicolor elements and no motion at all when color was mismatched, as if color mismatching vetoed the orientation-based motion.

We have thus shown that spatiotemporal matching of some image attributes (such as orientation in the above example) may elicit motion perception depending on whether some of the remaining attributes (such as color in the example) were matched coherently. We called these attributes veto attributes since their spatiotemporal mismatching entails the disruption of motion perception normally elicited by the matched attributes.[4,5] We have also shown that the strength of motion perception depends on the matched attribute; color, for example, was found to be a stronger motion token than orientation.

The totality of these results could be accounted for by the simple postulate that motion-detecting units extract relevant spatiotemporal information from subunits sharing the same attributes and preferences and characterized by similar filtering properties within a multidimensional sensory space. To illustrate, we note that this means that, if a motion-processing subunit is, say, achromatic, vertically oriented, and tuned to low spatial frequencies, the remaining subunits must display similar characteristics. Although this conclusion may appear trivial, it provides insight into what may be referred to as the spatial and temporal gluing of the lower-level visual units into some higher-order unit.[8] It should then be useful in understanding texture segregation to the extent that texture is defined in terms of either the global (or statistical)[9] or homogeneity[10] properties of a given (set of) attribute(s) characterizing a spatially distributed set of local features.

0740-3232/91/020386-08$05.00

As pointed out by Adelson and Bergen,[1] a motion-detecting unit can be looked on as a unit extracting spatiotemporal orientation and is therefore analogous to a x–y spatially oriented unit. Both units are hard wired. It is the specificity of these hard-wired connections that characterizes the so-called glue factor.

Spatially and spatiotemporally oriented units extract luminance (or chromatic) energy within a limited band of frequencies and orientations. This limited band can be defined spectrally through Fourier analysis. However, luminance and color are but two among several stimulus attributes that can define spatial (or spatiotemporal) patterns. For example, a set of identically oriented bars displayed in order to form an arbitrary pattern (say a large rectangle) immersed in a larger array of randomly oriented bars will pop out from the background. In fact the human observer will be able not only to provide a target/no-target type of response, but also to specify the shape or orientation of the composite target rectangle. This is a classical texture-type experiment whose principle can be generalized, not only to other lower-level image attributes such as color or size,[9–12] and to some of their conjunctions,[13,14] but also to supposedly higher-order attributes such as shape from shading.[15,16]

Recent studies have demonstrated that classical texture discrimination[9–12,17,18] can be accounted for by lower-level filtering through traditional size- and/or orientation-selective filters.[19–22] These studies emphasize the point that texture segregation may simply be based on the evaluation of the relative outputs of a homogenous population of filters at one or many scaling levels. Such an operation, however, does not lead directly to the specification of the shape of the discriminated array, or of the texture boundary. Voorhees and Poggio[20] proposed a nonparametric statistic design to compare local distributions and permitting the location of such boundaries. The biological substrate of this computation was not specified.

In contrast with this presumably time-consuming operation, a hard-wired operator, analogous to an orientation-tuned unit but at a higher processing level, would extract shape out of texture (or texture boundaries) instantaneously. Moreover, its neurophysiological substrate would be straightforward: It would consist of the spatial gluing of a set of multidimensionally tuned subunits. The specific question asked in the present paper is whether the rules governing this hard-wired gluing, i.e., the fine structure of a higher-order receptive field, are similar to those hypothesized by us as accounting for the veto effect in motion perception.[4–7] The present investigation is limited to the study of textures based on the coherent spatial grouping of color and orientation attributes.

STIMULUS RATIONALE

The stimuli shown in the left-hand column of Fig. 1 are close replicas of those used in our motion experiments.[3–7] Except for the two different shadings of the individual elements, which are meant to represent equiluminant red and green hues, they are scaled replicas of the stimuli actually used in the present experiments.

Although the individual texture elements (textels) are spatially displayed such that they can be grouped at will in vertical, horizontal, or ±45-deg rows, their grouping along the positive diagonal (left-hand column) is the only perceptually conspicuous configuration. It is due to the coherent grouping of color for stimuli A and B, of orientation for stimuli C and D, and of both color and orientation for stimulus E. Intellectually, grouping along the negatively sloped diagonals may also be achieved on the basis of orientation grouping (despite color alternation from one element to the

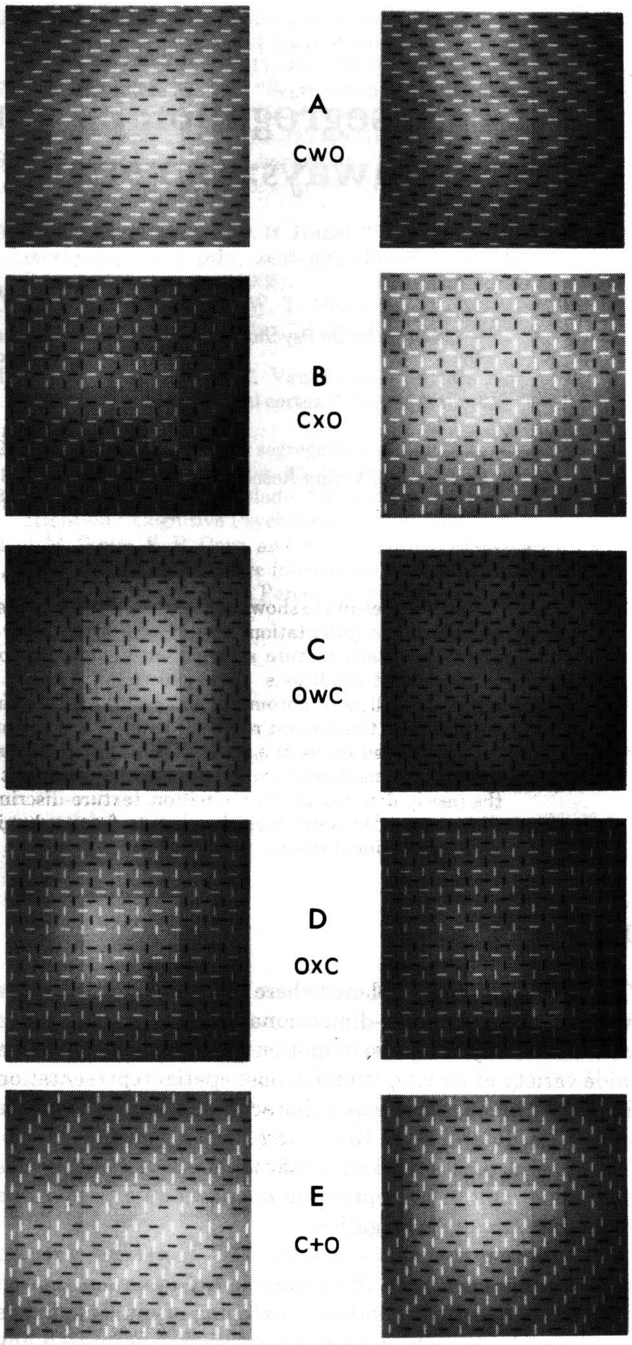

Fig. 1. Five stimulus pairs used in the present study: diagonal grouping (left-hand column) and chevron grouping (right-hand column). The two shadings of the bar elements represent red and green equiluminant hues. The background could be either dark or equiluminant yellow. A, Color grouping within orientation (CwO); B, color grouping across orientation (C × O); C, orientation grouping within color (OwC); D, orientation grouping across color (O × C); E, simultaneous grouping of color and orientation (C + O).

next; stimulus A, left-hand column) and on the basis of color grouping (despite orientation alternation; stimulus B, left-hand column). The empirical question is which of these possible groupings is also a valid perceptual grouping?

To introduce our terminology, notice that, for stimulus A, the two colors (shades of gray) are grouped in order to produce slanted stripes, while orientation is fixed for all textels. We refer to this condition as grouping color within orientation. By contrast, for stimulus B, color is grouped across orientation; this phrase is meant to denote that, while color is grouped to produce the same slanted stripes, the orientation of the textels alternates from 0 to 90 deg. In a dual manner orientation is grouped within and across color for stimuli C and D. Finally, stimulus E was obtained by grouping both attributes consistently. Hence stimulus configurations A and C will be referred to as grouping within conditions, and stimulus configurations B and D will be referred to as grouping across conditions.

Stimuli displayed in the right-hand column of Fig. 1 are identical with those displayed in the left-hand column in terms of the type of local grouping but different in terms of the resulting global spatial pattern (right-angle chevrons). In the experiments stimuli A–E were displayed by pairs of identical type of grouping but of different global configurations (i.e., diagonals and chevrons or arrow tails).

The design of the stimuli was meant to activate specific processing mechanisms. The rationale underlying the specification of these putative mechanisms has already been discussed in detail,[6,7] and it will be only briefly exposed here.

Discrimination performances of the members of any stimulus pair is necessarily based on the capacity of processing a specific type of spatial grouping (see also the Method section). The specific attributes used to obtain the five spatial-grouping types discussed above can be mapped onto a two-by-two sensory space. The two dimensions represent chromatic (C) and orientational sensitivity. They can be looked at as being binary to the extent that one includes C versus achromatic [i.e., luminance (L)] mechanisms, while the other includes oriented (O) versus nonoriented (nO) mechanisms. Within this sensory space it is easy to show that discrimination of stimuli A–E must be attributed to the unambiguous activation of the following mechanisms:

Stimulus	Activated Mechanisms		
A (CwO)	C–O	C–nO	
B (C × O)		C–nO	
C (OwC)	C–O		L–O
D (O × C)			L–O
E (C + O)	C–O	C–nO	L–O

Notice that none of these stimuli activates unambiguously the L–nO mechanism.

The notations between parentheses characterize the stimulus type according to the terminology discussed in Fig. 1. This stimulus decomposition leads to three interesting remarks.

First it permits the direct or indirect isolation of specific sensory mechanisms. Hence it provides information on their relative contribution to texture segregation. In view of previous studies concerned with form perception,[23] it is of theoretical importance to find out whether the mechanisms do contribute at all to texture segregation. Direct isolation can be obtained with stimuli B and D, which activate only one out of the four mechanisms in the given sensory space. Indirect isolation is achieved by comparing discrimination performances obtained with stimuli A and B, A and E, C and D, and C and E. For example, the comparison using A and B will isolate the relative contribution of the C–O mechanism.

Second it leads to qualitative predictions concerning the relative strength of texture discrimination. One expects that a texture whose spatial grouping activates a larger set of the various classes of mechanism described above will be better discriminated than a texture activating only a subset of these mechanisms. For example, stimulus E should induce the highest discrimination rates, stimulus A should induce higher discrimination rates than stimulus B, etc.

Third it leads to the rather peculiar prediction that, under equiluminant conditions, stimulus D (i.e., grouping of orientation across color) should induce zero or near-zero discrimination performances. Indeed, according to our analysis, texture discrimination for this stimulus depends on the activation of the L–O mechanisms only. At equiluminance, all L information is suppressed, the L mechanisms are thus silent, and the discrimination task should be impossible. More generally, all the discrimination tasks requiring the contribution of the L channels should be impaired under equiluminant conditions.

Most of the above predictions have been tested and verified in our motion experiments. If they are also verified for texture segregation, one should feel entitled to assume that motion and texture perception are but projections of each other onto two different physical spaces. A corollary of this assumption is that higher-order mechanisms subserving motion and texture perception extract spatial and spatiotemporal information in similar ways, namely, that they display similar anatomical connections.

All the experiments described below have been run under both nonequiluminant (dark background) and equiluminant (i.e., background equiluminant with the individual elements) conditions.

METHOD

Stimuli

The target stimuli (textels) were red and green bars oriented at 0 and 90 deg and displayed on a Sony Trinitron color video monitor (PVM-1270Q, P22 phosphor with a persistence of 62.5 μsec) driven by an ADAGE RDS3000 raster display under the control of a VAX 11/750 computer. The CIE x and y coordinates of the red and green guns of the monitor were (0.65, 0.31) and (0.29, 0.59), respectively (as measured with a Minolta Color Analyzer II, TV/2130). Stimuli were present on either a dark or yellow background obtained by the linear combination of the red and green hues set at equiluminance (see below). The L level of the stimuli was set near a mean of 39 cd/m^2.

The bars were 10.5 arcmin long and 1.2 arcmin wide (18 and 2 pixels, respectively). As shown in Fig. 1, the bars were displayed in 12 rows above and below the fixation point (not shown in Fig. 1), which was located in the middle of an empty area corresponding to a blank row. The distance between successive rows was 10.5 arcmin. Each row contained 13 bar elements 21 arcmin apart (measured at their midpoints) from one another. The whole display subtended 4.55 deg both vertically and horizontally at 200 cm from the observer.

Spatial grouping of a given attribute (orientation, color, or both) was used to form global patterns of either chevrons or (almost) continuous diagonals. These global diagonals could be positive (+45 deg, as shown on the left-hand side of Fig. 1) or negative (−45 deg), and the chevrons could point to the left or to the right (as shown on the right-hand side of Fig. 1). The stimuli were constructed in such a way that the transition from the bottom row of the upper half of the stimulus array to the top row of the lower half of the stimulus array was identical for the diagonal and chevron stimuli. This prevented their discrimination on the basis of local, transition clues. Avoiding such local clues would not have been possible with ±45-deg global diagonal pairs whose slant can be inferred from the matching of two adjacent textels only.

In all experiments the target stimuli were presented for a variable duration and were immediately followed by a mask stimulus built of yellow crosses presented at the same spatial locations as the target elements. The value of the yellow crosses was 82 cd/m^2. Unless specified otherwise, target duration was limited to 233 msec to prevent involuntary scanning behavior. The duration of the mask was fixed at 100 msec.

The red, green, and yellow (background) equiluminance points were set through the following flicker photometry procedure. The first step consisted of choosing an arbitrary green L level and adjusting the red L level in order to minimize flicker perception. This was done while using stimulus B in Fig. 1 and switching the element and background colors at a 30-Hz rate. The procedure was then repeated by fixing the experimentally found red equiluminance level and adjusting the L level of the green. Transitivity was verified. The second step consisted of adjusting the L levels of red and green such as to bring their sum (resulting in a yellow hue) to 40 cd/m^2 while keeping their ratio constant. This provided the yellow background. The third step was to apply once again the flicker-photometry technique with stimulus B in order to adjust the red and green hues, separately, for equiluminance with respect to the fixed yellow background. This completed the equiluminance setting procedure for the experimental conditions for which the background was dark.

In order to avoid, as much as possible, L-level clues owing to chromatic aberration under equiluminant background conditions, we randomized the red, green, and yellow L levels pixel by pixel as follows: (a) the L randomization range for yellow was set at 40 cd/m^2 ± 20%, (b) red and green equiluminant points with respect to minimum and maximum yellow values were adjusted through flicker photometry as above, (c) the whole stimulus array was rebuilt by randomizing each background pixel value within a ±20% range about the mean yellow value and each red and green pixel (belonging to the bar elements) within its minimal and maximal values as set through flicker photometry in step (b). The resulting average L levels (as measured photometrically) of the red, green, and yellow hues were the same as those set through flicker photometry under the condition for which the L level of each pixel was constant for a given hue in the constant pixel values conditions.

Each equiluminant point was the average of at least five adjustments per observer. The mean equiluminant values varied by less than 10% between observers, and their standard deviations were less than 2%.

Procedure

The two authors and a naïve female served as observers in all experiments. Discrimination performances for each stimulus pair (i.e., stimuli A–E in Fig. 1) were assessed as a function of stimulus duration. They were expressed as percentages-correct obtained by means of a two-alternative forced-choice procedure. The two configurations of a pair, each followed by the mask stimulus, were presented in random succession, and the observer was asked to decide which of the two temporal intervals contained the chevron configuration. The ± diagonal configurations and the leftward- or rightward-pointing chevron configurations were also chosen randomly from trial to trial. The use of the mask was meant to decrease performance since, in its absence, several stimuli produced 100% correct responses even for the shortest durations. The interval between the offset of the first mask and the onset of the second stimulus in a trial was set at 450 msec. Each performance assessment was based on 50 trials and repeated at least 3 times; thus each datum point presented in the Results section is based on a minimum of 150 trials. Both stimulus duration and stimulus types were randomized across sessions. Observers AG and GNP started with the dark-background conditions and observer TVP with the equiluminant-background conditions.

RESULTS

Figure 2 displays the discrimination performances of the three observers under dark-background conditions for the five stimulus pairs (see Fig. 1). The results display the following main trends for all observers.

As expected on grounds of temporal integration, there is a general trend of the discrimination performances to increase with stimulus duration for all stimulus configurations. For stimuli A, B, and E they reach 100% for durations in the range between 33 and 66 msec. For stimulus C, observers TVP and GNP attain better than 98% performances at 83 msec, while AG's performances seem to reach a 95% plateau at this same duration. For stimulus D, performances of observers AG and TVP continue to increase to the longest stimulus duration (233 msec), while still not reaching the 100% level; observer GNP reaches a plateau of approximately 76% correct for durations shorter than 133 msec. Of course, 100% performances would necessarily be attained if eye movements were made possible.

Stimulus-type related performances decrease in the following order: E, A ≈ B, C, D. Given our stimulus analysis (see Stimulus Rationale section) some of these order relationships were expected. Indeed, stimulus E is supposed to activate all four mechanisms in our hypothetical sensory space and is thus expected to induce the highest discrimination performance (E against A comparison averaged across observers; $X_1^2 = 55.9$, $p \ll 0.001$). $A > B$ and $C > D$ relationships were expected given that, in addition to the mechanisms activated by stimuli B and D, stimuli A and C also activate the C–O channel. However, the performance difference between A and B (averaged across observers) is not statistically significant, suggesting a weak or no contri-

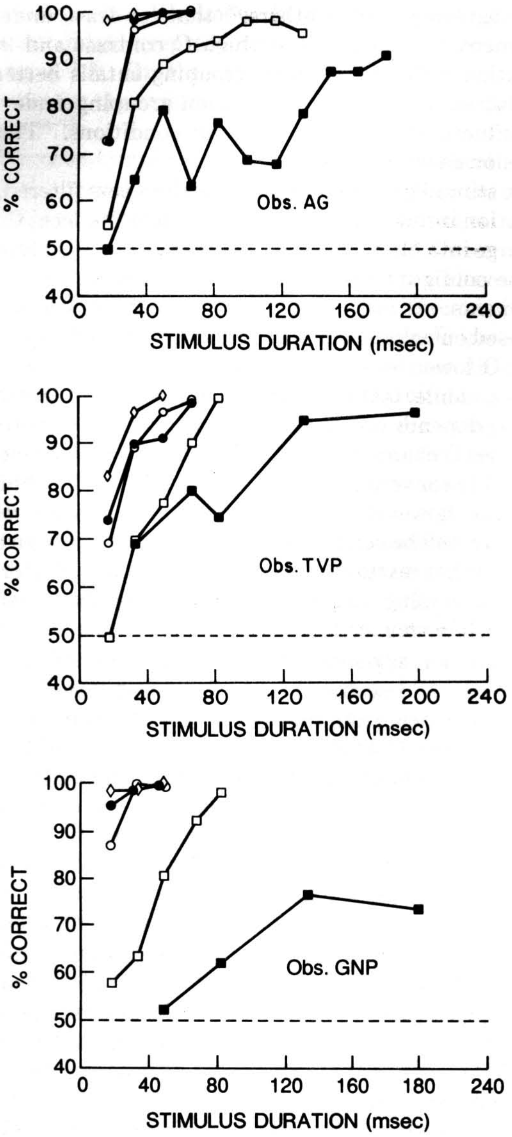

Fig. 2. Percentages of correct texture discrimination under dark-background conditions for the five stimulus pairs shown in Fig. 1 and for the three observers (AG, TVP, and GNP): open circles, stimulus A; filled circles, stimulus B; open squares, stimulus C; filled squares, stimulus D; diamonds; stimulus E. The dashed horizontal line shows the chance level.

bution at all of the C–O channel. In contrast, activation of the same mechanism produces a significant difference between C and D ($X_1^2 = 106.4$, $p \ll 0.001$). This discrepancy may reflect some characteristic inhibitory interactions between distinct channels (i.e., stronger inhibition between C–O and C–nO channels than between C–O and L–O channels). A simpler possibility is that the contribution of the C–O channel does not show up when added to performances that are already quite high (greater than 70%).

The $A \approx B > C > D$ rank order is identical to what we have obtained in the experiments on motion strength dependent on color and orientation matching over space–time (see Figs. 3 and 7 in Ref. 6; direction discrimination performances obtained with stimulus E were also the highest ones but were not reported in Ref. 6). Similar to those for the motion study, the present results indicate not only that C mechanisms can easily account for texture segregation, but also that color grouping (either within or across) is more effective in eliciting texture segregation than orientation grouping (within or across).

Figure 3 displays the discrimination performances of the three observers under equiluminant-background conditions for the five stimulus pairs. The consequence of eliminating all luminance clues is an overall decrease in texture segregation performances. For stimuli A, B, and E, this decrease is occurring essentially at stimulus durations equal to or shorter than 33 msec. For all observers, performances with these stimuli are at, or close to, the 100% level by 50-msec presentation. Performances obtained with stimulus C do not exceed 80% even for the longest duration for observers AG and TVP, while they are substantially lower for observer GNP. Increasing stimulus duration to 1683 msec (for which eye movements were unavoidable) produced a moderate increase in GNP's performances. By contrast, performances obtained with stimulus D are at chance level for all observers

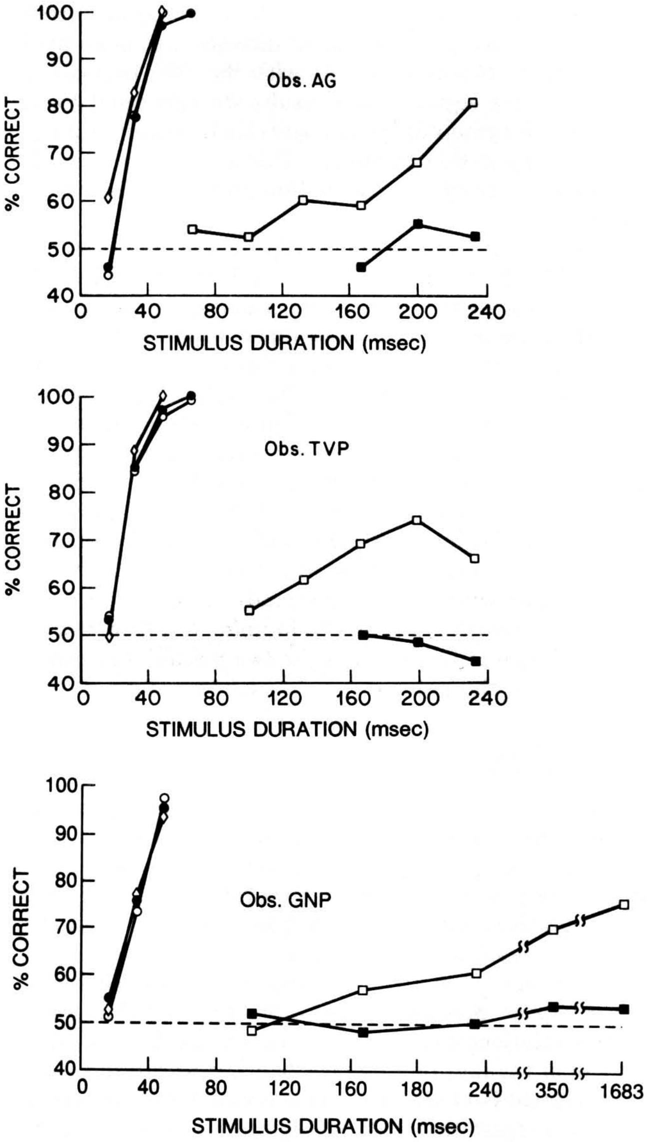

Fig. 3. Same as Fig. 2 but under equiluminant-background conditions.

and even for durations of as long as 1683 msec (for observer GNP). This last result is exactly the one predicted on the basis of our stimulus analysis (see Stimulus Rationale section): Under equiluminant conditions, stimulus D activates none of the four mechanisms within our hypothetical sensory space and cannot thus induce texture segregation. The same result was obtained in our series of motion studies.[3–7]

Similar to those of the dark-background conditions, stimulus-type related performances decrease in the following order: E, A ≈ B, C, D. The interpretation of this hierarchy is the same as that given above. While conditions C and D clearly entail lower performances than conditions E, A, and B, the differences among the latter three are quite small. The performance difference between A and B (averaged across observers) is not statistically significant, while the difference between condition E and conditions A and B is close to the 0.05 significance level ($X_1^2 = 2.9, 0.05 < p < 0.1$). Finally, condition C produces significantly higher performances than condition D ($X_1^2 = 105.4, p \ll 0.001$). Once again, these results suggest a weak contribution of the C–O channels when added to the contribution of the C–nO channels (stimulus A referenced to stimulus B) but a much stronger contribution when estimated directly (i.e., stimulus C referenced to stimulus D). As with the dark-background conditions, the equiluminant results strongly point to the fact that color grouping is a stronger clue for texture segregation than orientation grouping. This is, once again, consistent with our results in the motion discrimination experiments.[4,5]

An interesting comparison is that between performances obtained with stimulus D under dark-background conditions (filled squares in Fig. 2) and stimulus C under equiluminant-background conditions (open squares in Fig. 3). According to our analysis, these stimuli activate the L–O and C–O mechanisms, respectively. The results suggest that the former have a stronger contribution to texture segregation than the latter. Similar results were obtained in our motion experiments. The stronger contribution of the L–O mechanism (relative to the contribution of the C–O mechanism), however, may be due to the larger effective contrast of stimulus D (dark background) relative to the effective contrast of stimulus C (equiluminant background). Since we are not in a position to express C contrast in units of L contrast, the actual relative contribution of the two mechanisms cannot be assessed at this point.

DISCUSSION

Previous studies on visual search[24–31] have already demonstrated that color and orientation may produce perceptual grouping if adequately displayed. The stimuli used in the present experiments permitted, however, a systematic study of the putative underlying mechanisms and of their interactions as revealed by textural grouping.

Our results (1) support the theoretical analysis of the sensory space presumably activated by our stimuli (see Stimulus Rationale section), (2) confirm that the activation of three of the four putative mechanisms within this space (including the C–O and nO ones) may entail texture segregation, (3) suggest that texture segregation is governed by similar spatial-grouping (or gluing) rules as motion perception, and (4) show, in contrast with some previous reports[18] but in agreement with others,[25] that, at least under our experimental conditions (i.e., high C contrast and 90-deg orientation differences), color grouping entails better texture discrimination than orientation grouping under both nonequiluminant and equiluminant conditions. This last conclusion deserves a few comments.

If the stimuli presented in Fig. 1 are low-pass filtered, local orientation information will be lost, while the local C blobs will merge into blurred diagonals or chevrons, namely, exactly those configurations to be discriminated in the present experiments. It may then be argued that texture segregation based on color grouping is achieved through the activation of O lower-level C channels tuned to low spatial frequencies, while texture segregation based on orientation grouping depends on the activation of a whole population of lower-level O channels tuned to high spatial frequencies and presumably converging on a higher-order O unit. This distinction is, however, confusing because the latter mechanisms may well be color sensitive, too. The confusion is due to the fact that texture segregation is interpreted in terms of attribute grouping, rather than in terms of (putative) activated mechanisms.

Indeed, color, as a sensory dimension, is not orthogonal to orientation. As discussed in the Stimulus Rationale section, orientation-based grouping will necessarily activate either the L–O or the C–O channels or both. Conversely, color grouping may also activate the orientation channels (i.e., the C–O mechanisms). Texture segregation based on the exclusive (stimulus C under equiluminant conditions), or partial (stimuli A, C, and E) activation of the C–O channels would be either impossible or impaired if the stimuli were low-pass filtered. Thus, under some experimental conditions, the present results indicate that texture segregation must be achieved through the activation of localized C channels tuned to high spatial frequencies.[32] Although contested by some,[33] the existence of localized orientation-selective C channels has received recent psychophysical support from experiments showing that discrimination performance with targets of one orientation among distractors of a different orientation is high even under equiluminant conditions.[34]

The implications of the above considerations are twofold. The first (which is trivial) is that texture- (as well as direction-of-motion) discrimination performances should not be accounted for in terms of the matched physical attributes, but rather in terms of the sensory mechanisms activated by a given matching. This requires the specification of a sensory space relevant to the effective physical attributes of the stimulus. Such a sensory space is necessarily hypothetical and requires experimental support. The present experiments confirm the existence of three of the four postulated mechanisms, namely, the C–O, C–nO, and L–O channels, and by the same token demonstrate their contribution to texture segregation. The description of the sensory space, as hypothesized, is also supported by the veto phenomenon, namely, the impossibility of grouping orientation across different colors under equiluminant conditions. Experiments in progress show that luminance cannot play such a veto role with respect to orientation grouping and hence confirm our initial analysis based on the idea that the veto property depends on the parallel processing (i.e., through independent channels) of distinct values of a given attribute.[4,5]

The present stimuli do not permit the isolation of the L–

nO mechanism. We know, however, that achromatic-nO units are largely represented at different stages of visual pathways.[35,36] On the other hand, the neurophysiological evidence for the existence of C-oriented units is still under debate.[36–42]

The second implication of the above considerations concerns the neural processing subserving texture segregation. Under some stimulating conditions, texture discrimination may well be accounted for by the activation of lower-level, O, low-spatial-frequency C and/or L channels. Diagonal bars or chevrons will pop out in the low-pass filtered A, B, and E stimuli. Under some other stimulating conditions (stimuli C and D in the present experiments), low-pass filtering will deteriorate rather than improve discrimination performances.[32] Indeed, local orientation information, the only clue for grouping, will be lost. One must then assume that such local information is somehow glued together at some higher processing stage.

The gluing operation cannot be accounted for by lower-level filtering models such as those proposed by Bergen and Adelson[19] or by Vorhees and Poggio.[20] For the sake of generality, one would like to assume that the gluing process is also operational (but perhaps relatively less effective) under those conditions for which low-pass filtering could account for discrimination. Under this hypothesis, it is a matter of further research to find out which of the two processes effectively determines the discrimination performances.

If some of the present results are to be accounted for in terms of a gluing process, then one should admit that the receptive fields of the higher-order units are also oriented and that their orientation is independent of the orientation of their subunits. Thus a diagonal higher-order receptive field may be built of smaller receptive fields oriented vertically or horizontally.

Higher-order units of the type described above could be generalized to other visual primitives (e.g., local depth) whose spatial gluing cannot be explained in terms of early filtering. The only constraint to their construction would be that the input they receive comes from spatially displaced but otherwise identical lower-level units. The activation of these higher-order units would account for texture-boundary extraction without computation (in the sense of Vorhees and Poggio[20]). We are currently investigating the extent to which the discrimination between pop-out and non-pop-out conjunctions of attributes[13,14,25,30,43,44] may also be accounted for within such a theoretical framework.

ACKNOWLEDGMENTS

We thank Patrick Cavanagh and Bela Julesz for helpful discussions while this research was in progress. The paper has been significantly improved both in style and in content as a result of the constructive criticism of James P. Thomas and of an anonymous reviewer. A. Gorea was partially supported by grants from the Philippe Foundation and Fondation pour la Recherche Médicale.

REFERENCES AND NOTES

1. E. H. Adelson and J. R. Bergen, "Spatio-temporal energy models for the perception of motion," J. Opt. Soc. Am. A **2,** 284–299 (1985).
2. C. Chubb and G. Sperling, "Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception," J. Opt. Soc. Am. A **5,** 1986–2006 (1988).
3. A. Gorea and T. V. Papathomas, "Motion in color-nonoriented and luminance-oriented channels," J. Opt. Soc. Am. A **4,** P51(A) (1987).
4. A. Gorea and T. V. Papathomas, "Form and surface attributes in motion perception studied with a new class of stimuli: a basic asymmetry," Bell Laboratories Technical Memo 870924 (1987).
5. A. Gorea and T. V. Papathomas, "The concept of 'veto'-attributes in motion perception," Invest. Ophthalmol. Vis. Sci. Suppl. **29,** 265 (1988).
6. A. Gorea and T. V. Papathomas, "Motion processing by chromatic and achromatic visual pathways," J. Opt. Soc. Am. A **6,** 590–602 (1989).
7. T. V. Papathomas and A. Gorea, "Simultaneous motion perception along multiple attributes: a new class of stimuli," Behav. Res. Methods Instrum. **20,** 528–536 (1988).
8. As pointed out by an anonymous reviewer, the term gluing is traditionally used in reference to attentional processes. We use it here, however, to denote the combination of outputs from low-level units as inputs into some higher-order mechanism.
9. B. Julesz, "Texture perception," Nature (London) **290,** 91–97 (1981).
10. J. Beck, K. Prazdny, and A. Rosenfeld, "A theory of textural segmentation," in *Human and Machine Vision,* J. Beck, B. Hope, and A. Rosenfeld, eds. (Academic, New York, 1983).
11. B. Julesz, "Texture and visual perception," Sci. Am. **212,** 38–48 (1965).
12. J. Beck, "Perceptual grouping produced by changes in orientation and shape," Science **154,** 538–540 (1966).
13. K. Nakayama and G. H. Silverman, "Serial and parallel processing of visual feature conjunctions," Nature (London) **320,** 264–265 (1986).
14. D. Sagi, "The combination of spatial frequency and orientation is effortlessly perceived," Percept. Psychophys. **46,** 601–603 (1988).
15. V. S. Ramachandran and D. J. Plummer, "Preattentive perception of 3-D versus 2-D image features," Invest. Ophthalmol. Vis. Sci. Suppl. **30,** 160 (1989).
16. D. Kleffner and V. S. Ramachandran, "Perception of three-dimensional shape-from-shading," Invest. Ophthalmol. Vis. Sci. Suppl. **30,** 252 (1989).
17. B. Julesz, "Spatial nonlinearities in the instantaneous perception of textures with identical power spectra," Philos. Trans. R. Soc. London Ser. B **290,** 83–94 (1980).
18. J. Beck, "Textural segmentation," in *Organization and Representation in Perception,* J. Beck, ed. (Erlbaum, Hillsdale, N.J., 1982).
19. J. R. Bergen and E. H. Adelson, "Early vision and texture perception," Nature (London) **333,** 363–364 (1988).
20. H. Voorhees and T. Poggio, "Computing texture boundaries from images," Nature (London) **333,** 364–367 (1988).
21. J. Malik and P. Perona, "A computational model of human texture perception," Invest. Ophthalmol. Vis. Sci. Suppl. **30,** 161 (1989).
22. I. Fogel and D. Sagi, "Gabor filters as texture discriminators," Biol. Cybern. (to be published).
23. T. Troscianko, "Perception of random-dot symmetry and apparent movement at and near isoluminance," Vision Res. **27,** 547–554 (1987).
24. C. Bundesen and L. F. Pedersen, "Color segregation and visual search," Percept. Psychophys. **33,** 487–493 (1983).
25. S. Dehaene, "Discriminability and dimensionality effects in visual search for feature conjunctions: a functional pop-out," Percept. Psychophys. **46,** 72–80 (1989).
26. R. C. Carter, "Visual search with color," J. Exp. Psychol. Human Percept. Perform. **8,** 127–136 (1982).
27. E. W. Farmer and R. M. Taylor, "Visual search through color displays: effects of target-background similarity and background similarity," Percept. Psychophys. **27,** 267–272 (1980).
28. B. F. Green and L. K. Anderson, "Color coding in a visual search task," J. Exp. Psychol. **51,** 19–24 (1956).
29. G. Moraglia, K. P. Maloney, E. M. Fekete, and K. Al-Basi, "Visual search along the colour dimension," Can. J. Psychol. **43,** 1–12 (1989).

30. H. E. Egeth, R. A. Virzi, and H. Garbart, "Searching for conjunctively defined targets," J. Exp. Psychol. Human Percept. Perform. **10,** 32–39 (1984).
31. S. L. Smith, "Color coding and visual search," J. Exp. Psychol. **64,** 434–440 (1962).
32. To check whether orientation-based grouping could not be also due to low-spatial-frequency diagonal components in the image, we used both image-processing filtering techniques (luminance averaging across $N \times N$ pixels) and optical blurring with a Mylar screen to obtain low-pass filtered versions of the dark-background stimuli shown in Fig. 1. In particular, low-pass filtering of stimulus C on the left-hand side (i.e., O × C) to the point of not discriminating the orientations of the textels results in equal energy along both diagonals with no preferred perceptual grouping making discrimination of ±45-deg global diagonals impossible. The low-pass filtered image had the appearance of a checkerboard.
33. M. J. Nissen, L. L. Case, and L. Isenberg, "Perceptual independence of color and shape," Invest. Ophthalmol. Vis. Sci. Suppl. **28,** 362 (1987).
34. P. Cavanagh, M. Arguin, and A. Treisman, "Effect of surface medium on visual search for orientation and size features," J. Exp. Psychol. Human Percept. Perform. (to be published).
35. D. H. Hubel and T. N. Wiesel, "Functional architecture of macaque monkey visual cortex," Proc. R. Soc. London Ser. B **198,** 1–59 (1977).
36. S. M. Zeki, "The distribution of wavelength and orientation selective cells in different areas of monkey visual cortex," Proc. R. Soc. London Ser. B **217,** 449–470 (1983).
37. S. Shipp and S. M. Zeki, "Segregation of pathways leading from area V2 to areas V4 and V5 of macaque monkey visual cortex," Nature (London) **315,** 322–325 (1985).
38. D. H. Hubel and M. S. Livingstone, "Complex-unoriented cells in a subregion of primate area 18," Nature (London) **315,** 325–327 (1985).
39. M. S. Livingstone and D. H. Hubel, "Segregation of form, color, movement and depth: anatomy, physiology and perception," Science **240,** 740–749 (1988).
40. J. H. R. Maunsell and W. T. Newsome, "Visual processing in monkey extrastriate cortex," Annu. Rev. Neurosci. **10,** 363–401 (1987).
41. E. A. DeYoe and D. C. Van Essen, "Concurrent processing streams in monkey visual cortex," Trends Neurosci. **11,** 219–226 (1988).
42. C. R. Michael, "Laminar segregation of color cells in the monkey striate cortex," Vision Res. **25,** 415–423 (1985).
43. A. Treisman and G. Gelade, "A feature-integration theory of attention," Cognitive Psychol. **12,** 97–136 (1980).
44. J. M. Wolfe, K. R. Cave, and S. L. Franzel, "Guided search: an alternative to the feature integration model for visual search," J. Exp. Psychol. Human Percept. Perform. **15,** 419–433 (1989).