

## Short-range limitation on detection of feature differences

DOV SAGI<sup>1</sup> and BELA JULESZ<sup>2</sup>

<sup>1</sup>*The Weizmann Institute of Science, Rehovot, Israel* and <sup>2</sup>*AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA*

Received 15 January 1986; revised 10 June 1986; accepted 2 July 1986

**Abstract**—We studied the ability of observers to detect the presence of a clearly visible line segment against a background of line segments of different orientation. As we increase the number (density) of these background lines, we find that detectability does not behave monotonically. Adding a small number of background lines decreases detectability but if adjacent line segments are permitted to fall in close range, a further increase of background lines improves performance which eventually reaches a constant level. This suggests that detection of feature differences involves a short-range process. The range of this process is about two degrees or twice the length of the line segments used. Thus texture-gradients between different elements are only formed if the distance between these elements is not much larger than the average element size.

### INTRODUCTION

Recent theories of vision suggest that human vision operates in two modes, one is preattentive and the other is attentive. The first is assumed to be parallel in the sense that it can process visual information in many positions at the same time, while the other is assumed to be serial in the sense that the spatial extent of the stimuli that can be processed is limited (Treisman and Gelade 1980; Julesz 1981). Whereas the preattentive system is not limited by the area it can process, it is limited by the level of information it is capable of processing. Recently we have shown (Sagi and Julesz, 1985a), using a dense texture of line elements, that the preattentive system can locate feature gradients (local changes in feature densities) but cannot identify which features create the gradients. The ability of the preattentive system to detect only feature gradients suggests that this system is limited by short-range interactions. To investigate this possibility, we examined the ability of observers to detect a line target with known orientation, presented among other line segments of another orientation (see Fig. 1). This experiment is similar to the experiments conducted by Bergen and Julesz (1983) and in a more general way to those of Treisman and Gelade (1980). They both reported parallel processing when the target differed in some feature from the other items in the display. We were interested in determining whether the process that is responsible for parallel detection of feature differences can operate over long spatial range or whether it is limited to short range. To this end, we compared detectability of feature differences when the number of lines in the display was small, density was low and inter-line distance was large, with the detectability when the number of lines was large, density was high and inter-line distance was small. The inter-line distance could be constrained in order to explore processes that depend on the spatial separation between the lines.

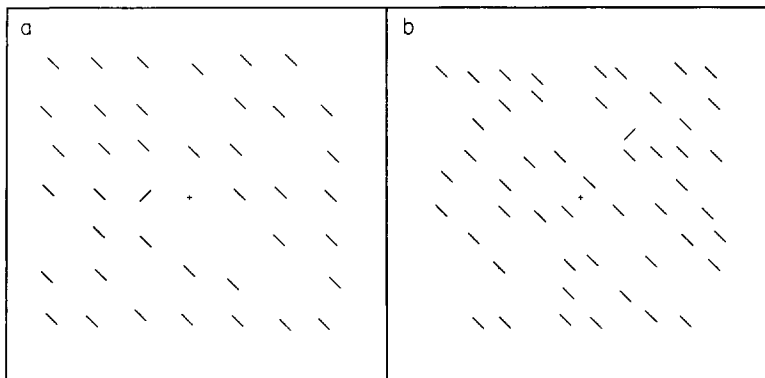
## METHODS

*Stimulus generation*

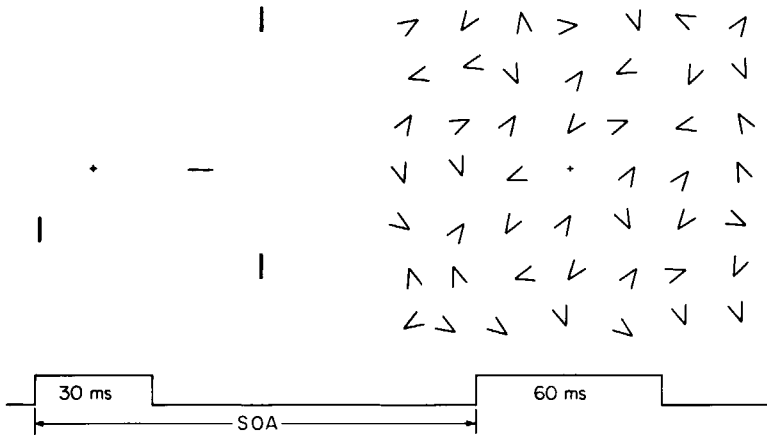
The stimuli were displayed on the face of a Hewlett Packard 1310 B cathode ray tube with P4 phosphor. The display was controlled by a PDP 11/23 computer and a special purpose hardware designed by W. J. Kropfl (described in Julesz *et al.*, 1976). The same PDP 11/23 controlled the experiments and collected the observers' responses from the keyboard. The keyboard was also used by the observers to initiate the stimulus display at their own pace. The stimuli consisted of an array of white line segments (1 deg length and 6 min arc wide) on a black background. The line arrangement is shown in Fig. 1. The square stimulus array was divided into 10 by 10 cells for the case of small minimal inter-line spacing or 7 by 7 cells for the large minimal inter-line spacing condition. Stimulus size was 20 by 20 deg of visual angle in the 10 by 10 cell case and 24 by 24 deg in the 7 by 7 cell case; minimal line separations (between line centers) were 2 and 3.4 deg respectively. In most experiments the number of lines presented was smaller than the number of cells with each cell having equal probability of being occupied by a line. The target lines were positioned randomly in a somewhat more restricted area, namely in cells of eccentricity larger than 3 deg but smaller than 8 deg of visual angle (in order to minimize effects of retinal inhomogeneity). The positions of the lines were randomized within an uncertainty area of 1 by 1 deg of visual angle. The observers were seated at a distance of 60 cm from the display.

*Stimulus presentation*

The stimuli were presented for 30 ms (see Fig. 2), thus preventing the possibility of more than one fixation during the exposure, though certainly visual persistence was longer. We limited the processing time available to the observer by masking the stimuli with patterns of randomly oriented V-shapes, presented for 60 ms at some time after the onset of the stimulus (this interval is called stimulus onset asynchrony, SOA). The SOAs were adjusted for each observer in order to obtain a detection rate of somewhat less than 100% correct for the easiest condition when the number of elements was two



**Figure 1.** Two of the displays used as stimuli: (a) 40 line segments in a 7 by 7 display, large minimal inter-line spacing. The observers had to detect the presence of the line having a different orientation from the others. (b) 50 line segments in a 10 by 10 display, small minimal inter-line spacing. Fixation cross in center is only presented before and after the stimulus and masker.



**Figure 2.** Temporal presentation of the stimulus. The stimulus (demonstrated by four lines) was presented for 30 ms followed by a blank period and then a masking frame (60 ms) containing randomly oriented V patterns.

(in order to be able to measure drop in performance for the somewhat more difficult conditions).

### *Psychophysical procedure*

In all experiments observers had to detect the presence of a single line of a known orientation among other lines on the display, of another known orientation. The number of lines (2 to 90) was varied randomly across blocks of 50 trials but was constant within blocks. In half the trials stimuli contained a line target (i.e. one of the lines had a different orientation) whereas in the remaining trials all lines had the same orientation. The orientation of the target and the background lines were kept constant within each block of trials though the positions of the background lines and the target were randomized from trial to trial. We used vertical targets among horizontal lines, vertical targets among diagonal lines and diagonal targets (45 deg) among diagonal lines, orthogonal to the target (135 deg). Blocks of trials with the same target and background orientations were grouped together.

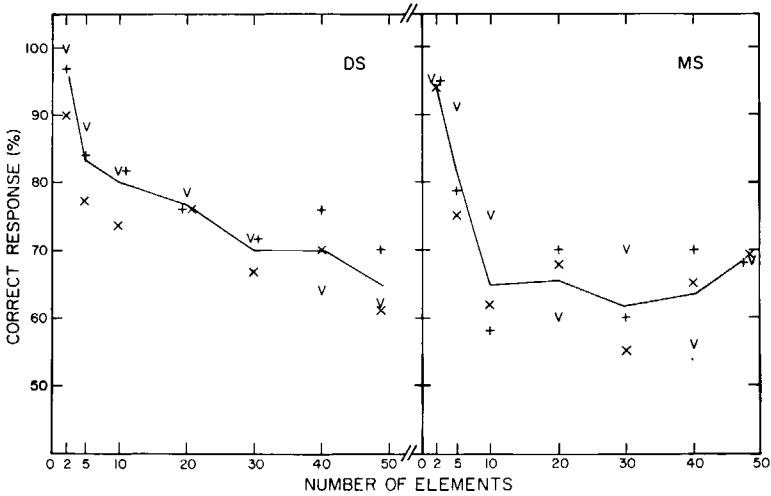
Detection rates were calculated as the average of the correct response rates of the two alternatives (target and no-target) in order to eliminate subjective preference toward one alternative or the other. Each session lasted 1–1.5 h.

### *Observers*

Four observers participated in these studies, all of them having normal or corrected to normal vision. One of the four observers was the first author (DS), while the others were unaware of the purpose of the experiments. All observers were well practiced in the experiments reported here; observers usually had lower performance in the initial phase of the experiments, but later they reached a constant level of performance. We did not take into account data from the earlier practice period.

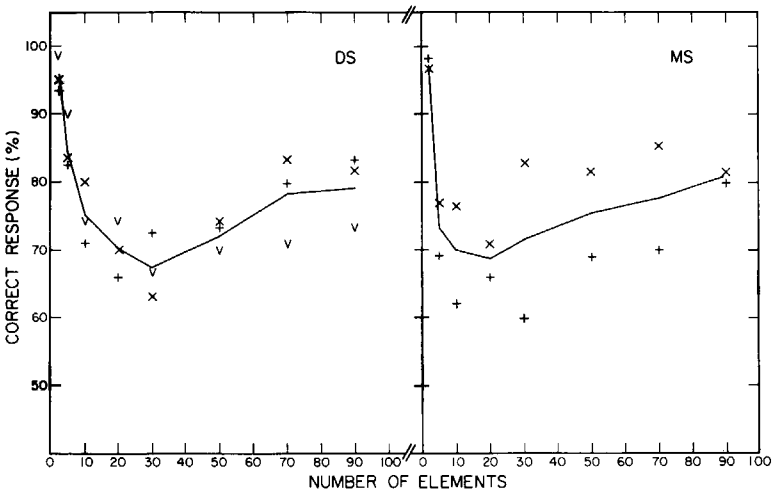
## **RESULTS**

Results for two different cases, representing the 7 by 7 cell array with large minimal inter-line spacing (3.4 deg) and the 10 by 10 cell array with small minimal inter-line



**Figure 3.** The dependence of performance on the number of lines presented for the large (3.4 deg) minimal inter-line spacing stimuli. Data of two observers, MS (SOA = 90 ms) and DS (SOA = 80 ms), for different orientations of target and background: (+) vertical target with horizontal background, (v) vertical target with diagonal background, (x) diagonal target (45 deg) and diagonal background (135 deg). Data points represent estimates based on 100 trials. Continuous curves represent average of the three conditions (300 trials per density point). Note the decline in performance as the number of lines increases.

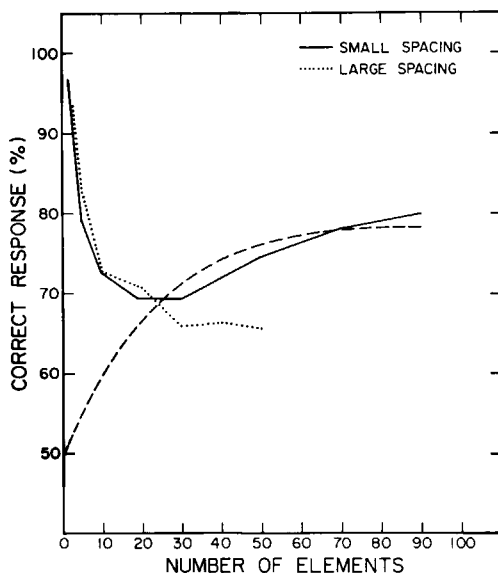
spacing (2 degs) are depicted in Figs 3 and 4 respectively. Different symbols represent different target and background orientations; the solid lines depict the averages of the different conditions. (Observer MS left before her 'V' data for Fig. 4 could be collected.) All the curves show a decline in performance as the number of lines in the display increases from two to twenty. However, the curves in Fig. 4 (corresponding to the 10 by



**Figure 4.** The same as Fig. 3, but for stimuli with small (2 deg) minimal inter-line spacing. Observers MS (SOA = 90 ms, 100 trials per data point) and DS (SOA = 60 ms, 175 trials per data point on the average). Note the improvement in performance when the number of background lines increases beyond 20 lines.

10 array) show an increase in performance when the number of lines is further increased, slowly reaching a second peak in performance in the  $N = 70$ – $100$  range (corresponding to 70–100% densities). Such an increase is not seen in the data from the large minimal spacing condition even with  $N = 48$  (98% density) as shown in Fig. 3, supporting the hypothesis that some short-range process is involved. Plotting the solid curve as an average of data points from different conditions is justified by the relatively small and seemingly unsystematic differences (except for observer MS in Fig. 4) in performance between the different conditions. There does not seem to be any qualitative difference between detection of a vertical line among horizontals (+), of a diagonal line among orthogonal diagonals ( $\times$ ), and of a vertical line among diagonals ( $\vee$ ). This is true despite the fact that the last case involves an orientation difference of only 45 deg, as opposed to a difference of 90 deg in the other two cases. We might expect a change in the system's behavior if orientation difference becomes smaller than some critical angle, but that angle seems to be smaller than 45 deg. Treisman (1985) used a similar detection task and measured reaction time. She found some performance asymmetries for small orientation differences between target lines and background lines; e.g., a tilted line could be detected in parallel among vertical lines but a vertical line was detected by a serial process when surrounded by tilted lines.

Note the different scales of Figs 3 and 4; both of them cover a range of line densities from 0 to 100%, but in Fig. 3 the range is up to 50 lines and in Fig. 4 up to 100 lines. When presented on the same scale, as in Fig. 5 (and taking the averages of the two observers over all line orientations of target and surround), the initial parts of the two curves coincide. The departure point is between 20 and 30 elements, beyond which the performance in the small spacing condition improves. This improvement can be



**Figure 5.** Data averaged from Figs 3 and 4, from the two observers. The curves of small (dotted curve) and large (continuous curve) minimal spacing overlap in the range  $N < 25$  indicating similar processes. For  $N > 25$  the continuous curve can be approximated by the dashed curve that is calculated on the assumption that discrimination is based on adjacent line comparison, within a distance of 2 deg of visual angle.

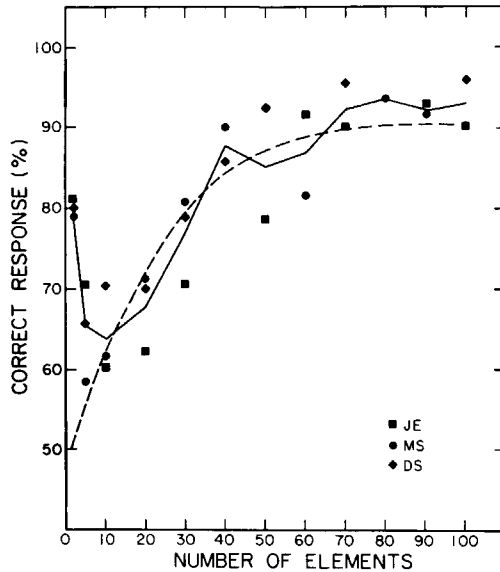
predicted on the basis of increased probability of having a background line adjacent to the target line as shown by the dashed line in Fig. 5. This curve is based on the assumption that discrimination performance can have some higher than chance value only if a background line is positioned in at least one of the four positions nearest to the target in order to determine a feature-gradient. The curve estimates the probability of this case with increased density. The theoretical maximum performance was used as a parameter to fit the curve to the data.

An interesting aspect of our results is the decline in performance as the number of lines increases from 2 to 30. This implies that with large distances between lines feature-gradients are too weak to be detected, only when element density is high are feature-gradients strong enough to be detected and performance becomes independent of the number of elements. The decline in performance with increasing number of elements, for a given SOA, implies serial processing (Bergen and Julesz, 1983); only when adjacent elements are within a critical distance (twice the length of line segments for the 10 by 10 cell case) that occurs above 70% density can feature-gradients be computed by a parallel preattentive system.

#### EXPERIMENTS WITH VARIABLE MASK DENSITY

Until now, we varied only the density of the stimulus arrays, but for masking used an array having 100% element density, that is, a constant masker. However, since we do not know much about the masking process, we might as well use a masker with the same number of elements as in the stimulus array. A variable masker might have some advantages, for example, since line density is varied we might expect some local interference effects due to the increased density (Bouma, 1970), effects that might change the observer sensitivity to the lines and thus cause changes in behavior with increased density. But, by changing the density of the masking elements (that are randomly oriented), we could change the amount of luminance (non-orientation selective) interference in the masking frame, in the same way they change in the stimulus frame. In a second set of experiments we presented mask elements only in positions having a line element in the stimulus display (first frame); thus the number of mask elements was varied along with the number of line elements. (As a result, we might expect an increase in local interference in the masking frame similar to the way these interferences increase in the stimulus frame, except for orientation dependent interferences since the masking elements are randomly oriented and the stimulus lines are not.)

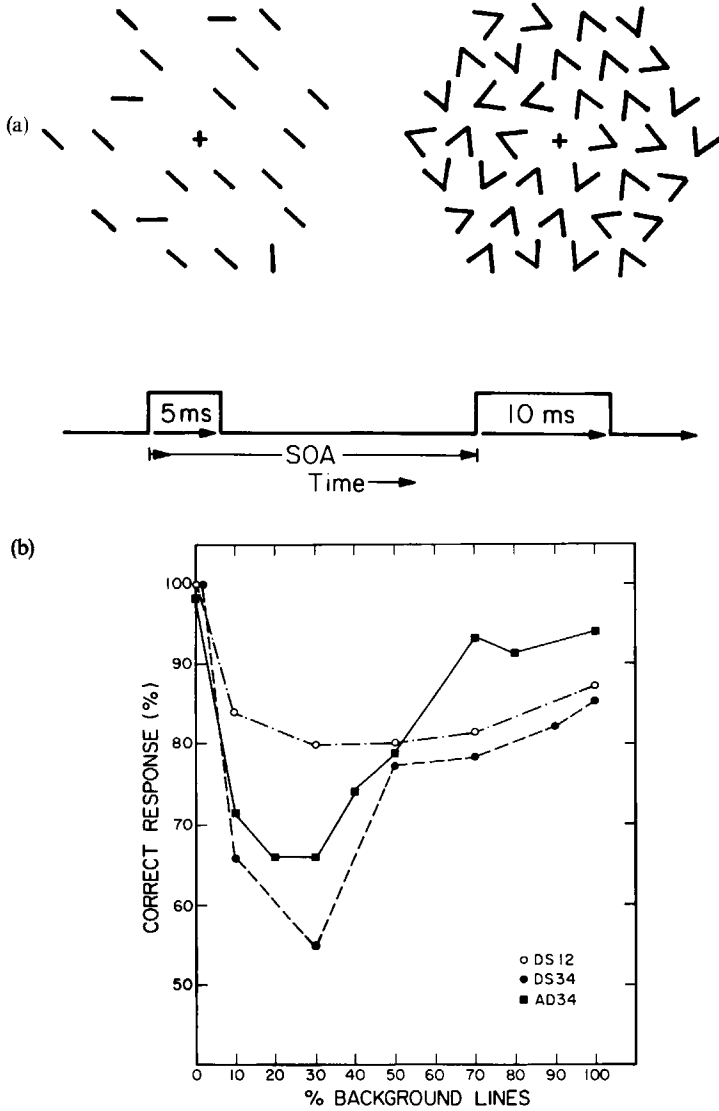
Figure 6 displays the results of one such experiment. In this case the stimuli consisted of the 10 by 10 array, having small minimal inter-element distance (2 deg). The results have the same characteristics as those of Fig. 4; an initial decline in performance followed by an increase reaching a plateau at the highest densities. Performance for more than 10 background lines can be predicted again on the assumption that target detection requires at least one neighboring line to form a feature-gradient of comparison between adjacent lines (dashed lines in Figs 5–6). Note that SOA is somewhat higher in these experiments. The increase in performance at high densities might be attributed to that increase in SOA, but the performance at low density is lower here than in the first experiment. Thus, individual mask elements seem to interact better with the individual lines, thereby creating a stronger mask and reducing performance. That is consistent also with the existence of local interferences within the masking frame.



**Figure 6.** Data from experiments where the number of mask elements was the same as the number of lines in the stimulus (three observers). The continuous curve represents the average of the three observers performance. The dashed curve is calculated on the assumption of comparison between adjacent lines within a distance of 2 deg, scaled to maximum performance of 90% correct. Observer JE (SOA = 100 ms) detected a vertical line among horizontal lines, observers MS (SOA = 120 ms) and DS (SOA = 100 ms) detected a diagonal line (45 deg) among diagonal (135 deg) lines. Data points are based on 150 trials on the average.

#### MORE THAN ONE TARGET: THE EFFECT OF DENSITY ON COUNTING

In all the experiments reported above we used a task of detecting one constant target among many background elements. In previous studies (Sagi and Julesz, 1985a, b) we have used a task in which observers had to detect different numbers of targets and discriminate in a two-alternative choice between 1 and 2 targets, 2 and 3, and 3 and 4 targets. We found then that performance was independent of the number of targets to be detected. Here we examined the effect of density change on such a task using the same experimental procedure and stimuli as in Sagi and Julesz (1985a, b) but with background densities as a variable (see Fig. 7a). The lines were arranged in a hexagonal pattern of 8 deg average diameter, targets were vertical and horizontal lines, embedded in a background of diagonal lines. All element positions had equal probability (between 0 to 1) of being occupied by a line, and this probability was kept constant through each block of trials. (As a consequence, the exact number of background lines in the different trials was not the same.) The masking stimulus contained all 36 elements, independent of the number of lines in the stimulus. Results for an experiment in which observers had to discriminate between 3 and 4 targets are depicted in Fig. 7b. In addition one observer's data for a 1 vs 2 targets discrimination are given. The curves have the same nonmonotonic behavior as the others (Figs 3–6) but interestingly the 3 vs 4 curves have a sharper dip than the 1 vs 2 curve and the detection curves in Fig. 4 (all cases with constant number of maskers). The change in shape implies that detecting four targets in low density background takes more time than detecting two targets, but if density is high enough (50% in this case) detection is independent of the number of targets and background density.



**Figure 7.** (a) Stimulus used in the counting task. The first frame (on the left) contained 3 or 4 vertical and horizontal line segments (the targets) and a variable number of diagonal lines (background). The second frame contained 36 randomly oriented V's (mask). In some experiments only 1 or 2 targets were presented. (b) Data from counting experiments, using display and targets as in (a). The abscissa represents the percentage of background lines presented, where 0% is the case where no background lines were presented and 100% is the case where all 36 lines were presented. Different symbols represent different observers and counting range e.g., observer AD (SOA = 150 ms, 50 trials per data point) counted from 3 to 4, observer DS (SOA = 90 ms, 100 trials per data point) counted from 1 to 2 and from 3 to 4.

Note that when the observers had to detect 3 or 4 vertical/horizontal targets, the addition of a few diagonal lines (in the case of 10% it is about 3 lines on the average) caused a dramatic reduction in performance from 99% correct to about 70% correct. Thus at the counting (or number discrimination) stage the line orientation is not known and there might be confusion between targets and background lines that can be

resolved by a slower process identifying the lines while counting them. Only at higher densities does performance not decrease when diagonal lines are added; on the contrary, adding more diagonal lines improves performance up to a constant high performance. Only at these high densities can the diagonal lines create a texture where the vertical/horizontal targets can 'pop out' and be counted as different from the relatively uniform background of diagonal lines. The transition point of 50% density can be regarded as the point where the background changes from uniform black (empty) to a uniform field of diagonal lines. This rationale might apply also to the data obtained from the other experiments described here (Figs 5–6).

To summarize, we find that observers' ability to detect feature differences depends on the number of elements presented when their density is below a critical value. This density corresponds to a critical distance between adjacent elements of about 2 deg, or twice the length of the line segments we used. Performance declines when the number of elements is small and density is low, but at high densities, performance improves and does not depend on the number of elements.

## DISCUSSION

We studied the effect of density and element number on detection of feature differences. Our results support the hypothesis that feature discrimination is not always performed in parallel. According to the hypothesis relating serial-parallel processes to attentive-preattentive processes, decline in performance with increasing element number indicates a serial (attentive) process whereas constant performance (independent of the number of elements) indicates a parallel (preattentive) process. (Townsend (1971) discusses the problems in identifying serial/parallel processes.) Given this assumption, our findings show that feature differences can be detected in parallel only when density is high and inter-element distance is short.

According to our findings, preattentive processing is limited to short-range interactions that operate only at a high density of elements; only at that high density can feature differences be computed in parallel. When density is low the preattentive system fails and a serial search is required. The operation of the preattentive system can then be accounted for by a simple system of feature detectors with local connections of inhibitory or excitatory type between similar detectors. One attractive mechanism is a generalized lateral inhibitory process between adjacent feature detectors that enhances feature gradients the same way luminance gradients are enhanced (Sagi and Hochstein, 1984, 1985). Similar interactions were suggested by Engle (1974, 1977) while studying visual conspicuity. He concluded that for non-homogeneous backgrounds the conspicuity of an object did mainly depend on restraining interactions from surrounding objects in the background and that interactions were at least size and luminance selective. He explained his results by assuming lateral inhibitory effects between channels for similar features.

Our results imply that detection of texture differences should be stronger at texture borders. The importance of texture borders was pointed out also by Mayhew and Frisby (1978). They observed that some textures differing in their spatial frequency composition could not be segregated when the border between them was covered by a thin luminance gradient. As Frisby concluded in his book (Frisby, 1980) "Segmentation by analysis of visual texture may sometimes depend crucially on detecting differences at the boundaries between regions, and when these are masked even by quite a thin line figure-ground perception via texture perception may fail".

Most electrophysiological studies concentrate on recording from single neurons responding to visual stimuli, but there is some evidence that single neuron responses are influenced by stimuli outside their classical receptive fields (Allman *et al.*, 1985). The spatial extent of motion interactions depends on the particular visual area examined and is greater for visual area MT than for area V2 in the monkey. Desimone *et al.* (1985), while recording from neurons in the V4 complex of macaque monkeys, found broad surround regions tuned to spatial frequency and color. These neurons would respond better to a grating limited to their receptive field than to a grating extending to their surround. If the same rule (surround selectivity) could be extended to orientation, we would have a possible neural implementation of our psychophysical results; only neurons having in their receptive field a line that differs in orientation from the surrounding lines would give a prominent response, all the other neurons that "see" in their receptive field and surrounds the same orientation would not respond. Such orientation-gradient sensitive units were recently found in V1 and V2 complex of macaque monkeys (DeYoe *et al.*, 1986). This suggested mechanism for the detection of feature differences is local, since the surround effect does not extend to the whole visual field and is capable of locating the targets to the extent of the localization accuracy of the responding neurons. It is suggested by these electrophysiological studies that localization accuracy depends on stimulus parameters and visual area, so we do not expect to find a single number for the range of this process.

Thus local changes in textures might be detected in parallel, whereas knowing what the features are that create this border might require some processes operating over longer spatial ranges, which might be serial and attentive (Sagi and Julesz, 1985a). It remains to be seen whether the range of this short-range process depends on the particular dimension on which detection is performed (for example, it might be longer for detection of motion differences), or on the particular stimulus parameters like element length, width, velocity, etc. We suggest that the dichotomy of vision based on short-range versus long-range processes (as in motion perception (Braddick, 1979)) has parallels to the dichotomy of preattentive versus attentive vision.

## REFERENCES

- Allman, J., Miezin, F. and McGuiness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanism for local-global comparisons in visual neurons. *Ann. Rev. Neurosci.* **8**, 407-430.
- Bergen, J. R. and Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature (Lond.)* **303**, 696-698.
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature* **226**, 177-178.
- Braddick, O. J. (1979). Low- and high-level processes in apparent motion. *Phil. Trans. R. Soc. Lond.* **B 290**, 137-151.
- Desimone, R., Schein, S. J., Moran, J. and Ungerleider, G. (1985). Contour, color and shape analysis beyond the striate cortex. *Vision Res.* **25**, 441-452.
- DeYoe, E., Knierim, J., Sagi, D., Julesz, B. and Van Essen, D. (1986). Single unit responses to static and dynamic texture patterns in macaque V2 and V1 cortex. *ARVO Abstracts 1986*.
- Engle, F. L. (1974). Visual conspicuity and selective background interference in eccentric vision. *Vision Res.* **14**, 459-471.
- Engle, F. L. (1977). Visual conspicuity, visual search and fixation tendencies of the eye. *Vision Res.* **17**, 95-108.
- Frisby, J. P. (1980). *Seeing*. Oxford University Press, Oxford.
- Julesz, B. (1981). Textons, the elements of texture perception and their interactions. *Nature (Lond.)* **290**, 91-97.
- Julesz, B., Breitmeyer, B. and Kropfl, W. (1976). Binocular-disparity-dependent upper-lower hemifield

- anisotropy and left-right hemifield isotropy as revealed by dynamic random-dot stereograms. *Perception* **5**, 129–141.
- Mayhew, J. E. W. and Frisby, J. P. (1978). Texture discrimination and Fourier analysis in human vision. *Nature* **275**, 438–439.
- Sagi, D. and Hochstein, S. (1984). The contrast dependence of spatial frequency channel interactions. *Vision Res.* **24**, 1357–1365.
- Sagi, D. and Hochstein, S. (1985). Lateral inhibition between spatially adjacent spatial frequency channels? *Percept. Psychophys.* **37**, 315–322.
- Sagi, D. and Julesz, B. (1985a). “Where” and “What” in vision. *Science* **228**, 1217–1219.
- Sagi, D. and Julesz, B. (1985b). Detection versus discrimination of visual orientation. *Perception*, **14**, 619–628.
- Townsend, J. T. (1971). A note on identifiability of parallel serial processes. *Percept. Psychophys.* **10**, 161–163.
- Treisman, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics and Image Processing* **31**, 156–177.
- Treisman, A. and Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychol.* **12**, 97–136.

## NOTE

1. We assume that the observers are guessing when the target line has no neighboring background lines, otherwise observers can detect orientation gradients with a probability  $p_i$  of correct responses,  $p_i$  may depend on SOA, acuity, line length, width, intensity, etc. Thus the total probability  $P$  of correct response is:

$$P = p_i(1 - p) + 0.5p$$

where  $p$  is the probability of having 4 empty neighbors. This formula is already corrected for guessing:

$$0.5 \leq p_i \leq 1 \text{ and } 0 \leq p \leq 1.$$

$p$  can be calculated exactly when the number of positions ( $N$ ) and the number of lines ( $n$ ) are given. The number of empty positions is  $N - n$ , and the probability of a given position being empty is  $(N - n)/N$ ; hence the probability of four positions being empty is approximately (for large  $N$ )  $p = ((N - n)/N)^4$ . Substituting  $p$  in the formula for  $P$  gives the theoretical curves plotted in Figs 5 and 6.