Vision Res. Vol. 30, No. 9, pp. 1377–1388, 1990
Printed in Great Britain.

0042-6989/90 $3.00 + 0.00

DETECTION OF AN ORIENTATION SINGULARITY IN GABOR TEXTURES: EFFECT OF SIGNAL DENSITY AND SPATIAL-FREQUENCY

DOV SAGI

Department of Applied Mathematics and Computer Science, The Weizmann Institute of Science, Rehovot 76100, Israel

(*Received* 27 *February* 1989; *in revised form* 17 *October* 1989)

Abstract—This work presents evidence for a second stage of spatial filtering in early vision. This second stage operates on the output of the well known linear spatial filters and integrates their thresholded responses with a center-surround weighting function. The evidence for the existence of this second stage comes from experiments where observers have to detect a Gabor signal with known parameters (target) among a varied number of other Gabor signals having orthogonal orientation (distractors). Detection performance on this task depends on the number of distractors and the distance between them: when the number of distractors is small performance deteriorates with an increasing number of distractors; however, when the number of distractors becomes larger performance improves with an increasing number of distractors. This improvement depends on the spatial-frequency of the signals and their spatial separation. Best performance is achieved when the spatial separation between signals is larger than three times their center wavelength but smaller than nine times their wavelength, implying a second stage filtering with a center size of six wavelengths and a total size of 18 wavelengths. This second stage of filtering may underlie our ability to detect certain texture boundaries preatentively.

Visual search   Parallel processing   Spatial filters   Attention   Texture

INTRODUCTION

Early visual processing stages seem to involve spatial-temporal filtering by a variety of linear filters. It has been suggested that their responses produce a compact representation of the input image (Daugman, 1985; Watson, 1983). These filters are sensitive to luminance variation across space on different scales, thus occupying a limited area in the spatial-frequency domain as well as in retinal space. Our ability to model the visual system that way stems from contrast detection experiments (Campbell & Robson, 1968; Graham & Nachmias, 1971; Watson & Robson, 1981; Wilson & Bergen 1979), from contrast discrimination experiments (Sagi & Hochstein, 1983) and from masking experiments (Daugman, 1984; Legge, 1979; Stromeyer & Julesz, 1972). It is assumed that these filters (channels) cover the whole visual field and thus process it in parallel (Graham, 1989; Watson, 1983; Wilson, 1983). However, it is not clear whether the information represented in this parallel system can be used simultaneously by a higher level decision stage. For example, can we compute phase relationships or orientation simultaneously across the visual field, taking advantage of the large number of filters responding to different locations in the visual field? Models of detection assume that all channels can be monitored simultaneously and thus incorporate an assumption of statistical (probability) summation across space (Graham & Robson, 1987; Wilson & Bergen, 1979) and spatial-frequency (Graham & Nachmias, 1971). According to this assumption, channel activity (whether below threshold or above) is monitored across the whole visual field by a decision stage, thus predicting improved detection rate with increasing stimulus area (Graham & Robson, 1987). The individual channels that produce above threshold responses, do not have to be identified, since it is sufficient for the decision stage to know that one of the many channels crossed its threshold. However, channels having different sensitivity curves can be labeled according to their most sensitive orientations and spatial-frequencies. Watson and Robson (1981) asked whether the label of an individual channel (orientation or spatial-frequency) can be identified at contrast detection threshold. They found, using localized stimuli, that patterns can be identified at their detection threshold, thus confirming the labeling hypothesis. In this study we examine the ability of the visual system to carry out many identification tasks at different locations in the visual field. Considering the results from search experiments (Bergen & Julesz, 1983a; Treisman & Gelade, 1980; Sagi & Julesz, 1985) we may expect a failure of parallelism for some visual tasks.

In a search experiment observers are confronted with the task of detecting the presence of a target when the visual field is cluttered with a variable number of distractors. Since target position is not known, all displayed items (target and distractors) have to be processed at least to the degree of being discriminable (as pointed out by Duncan, 1985). The interesting variable is usually detection time, as measured by reaction time (Treisman & Gelade, 1980) or by error rate when stimulus processing time is limited by backward masking (Bergen & Julesz, 1983a; and see Methods section below). In general, detection time increases with increasing number of distractors unless the target differs from the distractors by a simple feature (orientation, size, color etc). In this latter case search is said to be carried out in parallel across the visual field, or preattentively. When a target differs from the distractors in the way their features are combined (spatial-relations between line segments as in T vs L, or conjunctions of features), preattentive detection fails and an attentive serial search is observed (Bergen & Julesz, 1983a; Treisman & Gelade, 1980). These findings raise two interesting issues: (a) what are the basic elements (features, channels) of preattentive vision? Do they agree with the known spatial channels? (b) what is the mechanism that enables parallel selection of a target embedded in a field of distractors?

As for the first issue, the list of preattentive features (textons) includes orientation, color, size, direction of motion and eye-disparity (Bergen & Julesz, 1983b; Dick et al., 1987; Nakayama & Silverman, 1986; Treisman & Gelade, 1980). Some other spatial features as line-crossings (Bergen & Julesz, 1983a) and terminators (Julesz, 1981; Treisman, 1985) are less obvious (Gurnsey & Browse, 1987; Kröse, 1987). A major problem here is the absence of a quantitative definition of features. Alternatively, the spatial filtering approach offers a measurable set of "features" that is general enough to be used for modeling human performance (Beck, Sutter & Ivry, 1987; Bergen & Adelson, 1988; Caelli & Moraglia, 1985; Daugman, 1987; Fogel & Sagi, 1989; Landy & Bergen, 1989; Malik & Perona, 1990; Nothdurft, 1988, 1990; Rubenstein & Sagi, 1990; Turner, 1986; Voorhees & Poggio, 1988; but see Julesz & Kröse, 1988). However, in its simplest form (discrimination is based on differences between averaged responses) the spatial filtering approach cannot account for some interesting cases such as terminator based discrimination (Fogel & Sagi, 1989). It was shown that this problem can be resolved by considering the variations of filter responses across space and orientation, taking into account the combined filter selectivity for space/orientation/spatial-frequency (Rubenstein & Sagi, 1990). Such a combined selectivity is consistent with the findings of Caelli and Moraglia (1985) who found enhancement in texture segregation tasks when figure differed from ground by both spatial-frequency and orientation.

The strong interdependence between spatial-frequency and orientation is a basic property of spatial filters. More than that, these two dimensions seem to be inseparable (Daugman, 1984). Thus, if spatial-frequency/orientation channels are available for preattentive vision, we may expect preattentive detection of a target that differs from distractors in the way spatial-frequency and orientation are combined. On the other hand, it is possible that preattentive vision operates at some higher processing level where orientation and spatial-frequency are encoded by separate units and thus their combinations are not readily accessible (Treisman, 1985; Walters, Biederman & Weisstein, 1983). Experimental evidence (Sagi, 1988) seems to favor the first option, thus indicating that preattentive vision operates at a processing level where spatial-frequency and orientation are not encoded separately.

The second issue concerns the mechanism that enables parallel detection of a target embedded in an array of distractor. One way to solve this problem is by postulating a processing stage where different features are being represented separately (feature maps) and the detection of a target having a unique feature is performed by looking for *total* (global) activity at the corresponding map (Treisman, 1985). If distractors and targets activate different maps, detecting activity in the target map is not influenced by the presence of distractors and their number. Thus preattentive vision is assumed to operate in feature space (Barlow, 1981) without having access to spatial information

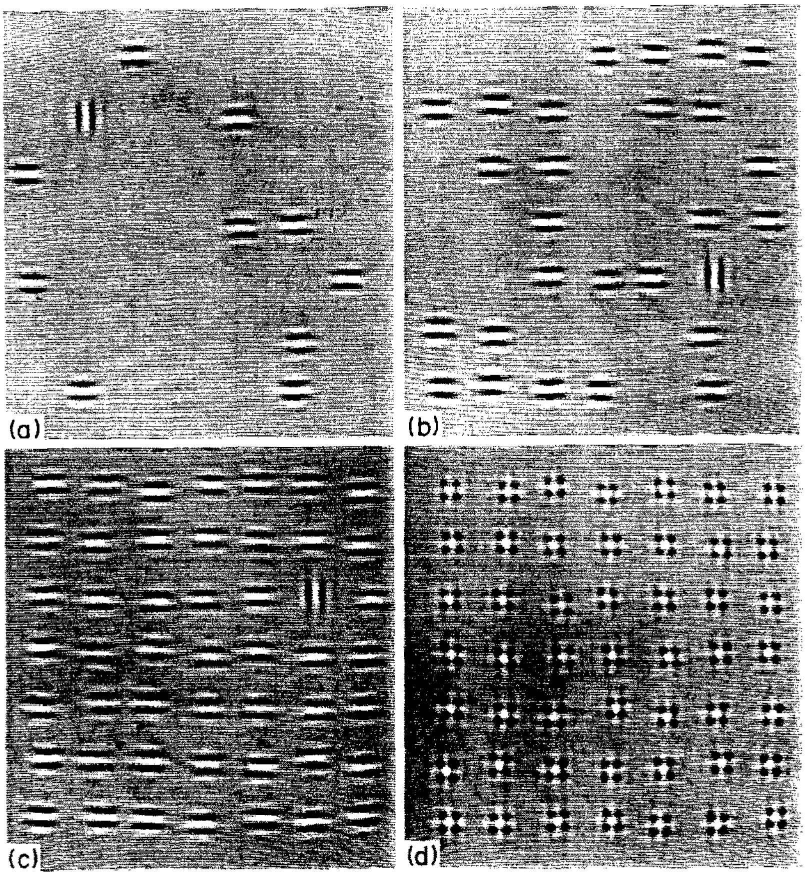

Fig. 1. Typical stimuli used in the experiments using 7 × 7 displays with different signal (distractor) densities: 12% (a), 50% (b) and 100% in (c). The mask frames (d) had always 100% density. The observers had to detect the presence of the vertical patch in (a), (b) or (c). In the 3 × 3 displays only the even rows and columns were used, thus doubling the inter-signal spacing. Stimulus contrast is enhanced here for demonstration clarity.

or to specific locations (Treisman & Gelade, 1980).

Another way to model the preattentive system is by assuming a mechanism that detects *local* differences in some feature maps where targets and distractors are mapped according to their spatial locations (Fogel & Sagi, 1989; Landy & Bergen, 1989). In this case, preattentive detection is possible only if the target generates a strong local difference signal in comparison with local difference signals generated by the distractors (Rubenstein & Sagi, 1990; Sagi, 1988). It was found that local differences generated by targets of different orientations cannot be discriminated preattentively (Sagi & Julesz, 1985, 1987; Braun & Sagi, 1990). For example, the orientation of two Gabor signals (presented in different regions of an otherwise empty visual field) cannot be identified simultaneously (Braun & Sagi, 1990) and there are also experimental results which show a slow down in detection of local color differences in the presence of shape differences (and vice versa, Pashler, 1988; Wertheim, 1981). On the other hand, shape and color local differences can be detected in parallel (Arguin & Cavanagh, 1988) and local feature differences can be localized preattentively (Sagi & Julesz, 1985). Thus we can assume that all local difference signals are combined according to their spatial location, regardless of their label (orientation, color etc.). This scheme relies heavily on experimental results showing strong dependence of preattentive detection on inter element (targets and distractors) distance. Detection of feature differences (orientation of line segments) was found to improve with increasing elements density, both in texture discrimination tasks (Nothdurft, 1985) and in search tasks (Sagi & Julesz 1987). The experiments reported here extend these latter findings to stimuli consisting of elements that are limited in their spatial-frequency content.

The aim of this work is to generate some data base that will make it possible to model the somewhat higher level phenomena (detection of feature differences) that are revealed in the texture and search experiments, making use of known low-level properties of the visual system. Accordingly, the experiments described here are very similar to those performed by Sagi and Julesz (1987) with the main difference of using patterns which are band-limited in the spatial-frequency domain and the space domain, thus probably stimulating a small number of spatial filters (channels). The visual task used is the detection of a vertical target in the presence of horizontal distractors (serving as noise). Typical stimuli are depicted in Fig. 1(a–c) demonstrating different distractors' densities. The basic phenomenon explored here is the improvement in target detection as distractors' density is increased. The use of similar experimental design comes as an attempt to bridge a gap between search experiments and standard contrast detection/discrimination experiments.

The main finding of Sagi and Julesz (1987) is that detection of the presence of a line segment target among line segment distractors of orthogonal orientation is not monotonically related to the number of distractors. When the number of distractors is small and the distance between any two elements in the visual field is large, performance declines with increasing number of distractors; however, above some critical distractor number (density) performance improves with increased number of distractors. A reasonable assumption is that only the latter phase relates to parallel processes engaged in the detection of the target, while the first phase (few distractors, large distances) relates to a serial scan over the isolated distractors and target. While in this article the main emphasis is on the parallel stage, there is accumulating evidence for serial processing in the low density case (Braun & Sagi, 1990). The experiments reported here explore a wider range of constraints on the parallel phase, mainly its dependence on the frequency content of the patterns distributed across space and on their spatial configuration. The result is a more comprehensive description of the mechanism underlying detection of feature differences (over short range).

The theoretical framework adopted here is the one used by Fogel and Sagi (1989) for modeling texture segregation processes. The model has two filtering stages. Filters of the first stage are standard spatial-frequency/orientation selective filters and those of the second stage are spatial-frequency but not orientation selective. A second stage filter, termed here hyper-filter, receives input from filters having the same parameters except for spatial location and phase. (This latter assumption accounts for the phase insensitivity of preattentive vision, Rentschler & Treutwein, 1985.) A hyper-filter can be viewed as a Laplacian–Gaussian linear operator operating on the squared (or thresholded) output of the first stage linear filters. Taking the difference of Gaussian (DOG) approximation to the Laplacian of Gaussian, a hyper-filter can be be characterized by two space parameters: excitatory and inhibitory (reflecting the standard deviations of the two Gaussians). The first parameter specifies the range of integration of the filter's output (amount of spatial blur introduced into the filter response map) and the second specifies the range over which feature differences can be detected by a hyper-filter. It is assumed that detection of feature-gradients is mediated through hyper-filters. In cases of uniform textures, hyper-filters will show weak responses that may correspond to texture local variability (attenuated by hyper-filter blur), however, the introduction of a different feature (or feature singularity) will yield a strong hyper-filter response. In cases of sparse textures (large inter-elements spacing), responses to texture local variations (between texture elements and background luminance) may compete with responses to feature singularities. In search tasks this may occur when only a small number of distractors are present and the inter-distractor distance is larger than the inhibitory space constant of the hyper-filter. Parallel target detection may then fail and attentive search may be necessary to resolve the hyper-filter response ambiguity. In addition, the hyper-filter response level may depend on how much of the hyper-filter integration area is covered by the target. Thus, for the task used here, detection of a vertical Gabor signal in the presence of horizontal Gabor signals, detection rate may depend on target size relative to its hyper-filter summation area and on signal density.

METHODS

Stimulus generation

The stimuli (see Fig. 1) were displayed on the face of a Conrac video monitor, with an average luminance of 100 cd/m^2, using an Imaging Technology frame buffer with a 256×256 pixels resolution at a frame rate of 50 Hz (noninterlaced). Stimulus generation and display was controlled by a SUN-2 work-station, using a special frame-buffer driver for stimulus timing control. The stimuli consisted of an array of Gabor patches, each patch occupying 32×32 pixels, where the Gaussian envelope had a spread of 16 pixels between $1/e$ points (see Fig. 1). The Gaussian envelop was modulated by a cosine function with a variable wave length between 4 and 32 pixels. The stimulus array was divided into 7×7 cells or 3×3 cells depending on the experiment, in order to control the minimal distance between the patches. Accordingly, the number of patches presented in each stimulus varied between 1 and 49 or 1 and 9. The position of the patches was randomized around each grid position within a range of 6 pixels.

We used two viewing distances of 180 and 80 cm, thus the stimulus occupied an area of 8×8 of 18×18 deg respectively. The Gabor patch size then was 1 or 2.3 deg and the inter patch separation was 1 (7×7 grid) or 2 deg (3×3 grid) at the longer viewing distance and 2.3 (7×7 grid) or 4.6 deg (3×3 grid) at the shorter viewing distance. Targets (to be defined later) were positioned in somewhat more restricted area, in order to keep their distance from fixation point near constant. They were placed on the sides of a square with dimensions of 4×4 (large viewing distance) or 9.2×9.2 deg (small viewing distance) centered at the fixation point.

Stimulus presentation

The stimuli were presented for 40 msec (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 a full screen of Gabor patches (7×7 or 3×3 as defined above), each of them generated as a superposition of the target patch and the distractor patch with a total contrast of 100% and 100 msec presentation time. In addition, the relative position (phase) between mask element and stimulus elements was randomized (due to jitter in positioning the elements). As a result of the high mask energy (contrast and duration) and its inconsistent relationship with the stimulus, the task of detecting the target proved to be impossible at some small SOA (Stimulus Onset Asynchrony). At these SOAs stimulus and mask were probably superimposed perceptually by visual persistence.

Target and distractors had the same contrast, ranging from 15 to 60% (defined as the ratio of

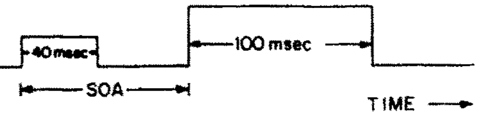

Fig. 2. The temporal sequence of stimulus and mask presentation on each trial. The display luminance was kept constant, stimulus and mask were defined by contrast where the mask contrast always 100% and the stimulus contrast was adjusted for each observer and spatial-frequency separately.

the cosine amplitude to the average luminance, thus reflecting the actual contrast only at the center of the Gabor patch). The relatively large range of contrasts used reflects the dependence of observers' sensitivity on spatial-frequency. Contrast was adjusted so that observers' performance at the lowest density used (a single target or distractor) was around 90% correct at SOA of 60–80 msec.

Screen luminance did not deviate from the expected linear relationship to the input signal by more than 5%. Stimulus patterns were digitized using 256 gray levels.

Psychophysical procedure

In all experiments the observer had to detect the presence of a Gabor patch with vertical orientation (target) among other patches with horizontal orientation (distractors). Between each block of 50 trials the average number of distractors was varied and within each block the probability of each grid position having a distractor (distractor density) was kept constant: 0, 0.12, 0.25, 0.5, 0.75 or 1.0, with the exception that for zero density one distractor was presented in trials without a target. A target was presented on about half of the trials (probability of 0.5 per trial). The observer had to respond 0 for target absence or 1 for the presence of a target, using the computer terminal keyboard. 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 which in any case was not significant). Each session lasted 1–1.5 hr.

Observers

Four observers participated in these experiments, all of them having normal or corrected to normal vision. One of the four observers was the 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. Only the latter phase was used in the data analysis. Observers were tested on 100–500 trials for each data point.

RESULTS

The first result of interest here is the dependence of target detection rate on the density (or number) of distractors. Results for the 7 × 7 cell array with small minimal inter-patch spacing for far and close viewing conditions are depicted in Figs 3 and 4 respectively, where in each figure data for three different spatial-frequencies are presented. The first point to note is that all curves show a decline in performance as the number of distractors increases, this initial decline can be from almost perfect performance to close to chance performance in some cases. However, some curves show an increase in performance when the distractor density is further increased, slowly reaching a maximal performance level at 100% distractor density. This maximal performance is sometimes equal to the performance at zero density.

It can be seen that curves showing an increase in performance as signal density increases are obtained from experiments using high spatial frequency Gabor signals. The transition seems to be around 2 c/deg in our case. We wanted to see whether the critical frequency for high density increase depends on other spatial parameters of the image as the minimal spacing between patches. Thus we ran experiments using

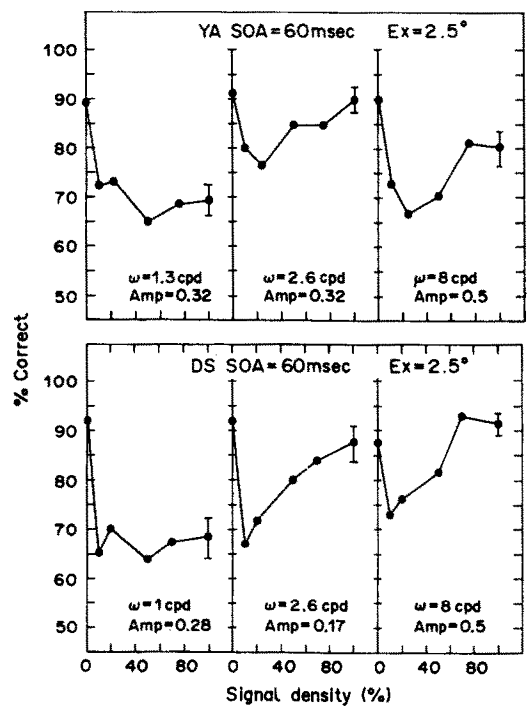

Fig. 3. Dependence of detection performance on patch density at three different spatial-frequencies (two observers). Stimulus size is 8 × 8 deg and target eccentricity is 2.5 deg on average. At 100% signal density the stimulus contained 49 signals arranged on a 7 × 7 grid with 1 deg spacing. The error bars represent averaged standard errors. Note the difference between the low-frequency and the high-frequency curves.

sparser grids having only 3 × 3 locations and varied the number of distractors from 1 to 9 as the density was changed from 0 to 100%. Results are depicted in Fig. 5. The main point to note is that performance, on high spatial frequency stimuli, does not show a strong improvement as signal density is increased. This latter improvement is still evident in the medium frequency data. On the other hand, the low frequency data show some increase in performance when signal density is increased. This later result demonstrates that large inter-element spacings improves parallel detection of low-frequency Gabor signals.

In order to quantify the "efficiency" of parallel detection in our experiments, we can define an arbitrary measure taking into account the increase in performance toward higher densities. One way to describe this increase is to take the difference between performance at 100% density and the minimal performance (over all densities) and to scale it with the initial decrease from 0% density to the mimimal performance point. Thus $PR = (P_{100} - P_{min})/(P_0 - P_{min})$. $PR = 0$ means that the performance at high density is equal to the minimal performance, thus there is no high density increase, while $PR = 1$ means that the high density increase is equal to the low density decrease. In general, PR can be larger than one but this is rarely observed in the data. The case $P_0 = P_{min}$ is never observed in the present experiments so practically PR is quite well defined. The dependence of PR on spatial frequency is depicted in Fig. 6 for two separations and two eccentricities. The graphs confirm the observation made above: for the small patch separation parallel detection improves with increasing the patch spatial frequency, thus producing best performance in the high spatial-frequency range, while at the larger separation best performance is obtained for medium spatial-frequencies. At these medium spatial-frequencies the large separation PR is larger than the small separation PR. Thus reducing patch separation can reduce the efficiency of parallel detection. This phenomenon may be a result of lateral masking between the target and the distractors, and is expected if filter size is larger than signal size.

It should be pointed out that the patch size (controlled by its Gaussian standard deviation) is the same for small and large separations and that separation is measured between center of patches, thus effective separation is probably smaller. This bias is larger in the small separation case where the Gaussian spread is half the

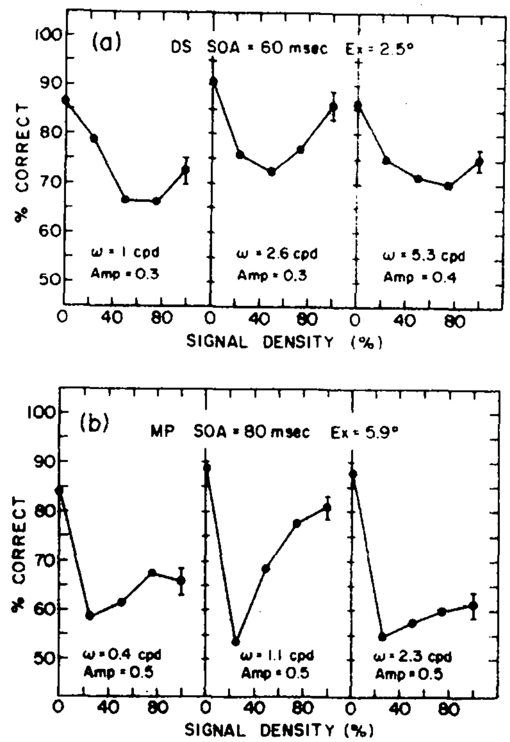

Fig. 5. The same as Fig. 3 (a) and Fig. 4 (b) but using a 3 × 3 patch stimuli with 2 deg (a) or 4.6 deg (b) patch spacing. Note the change in the high-frequency region.

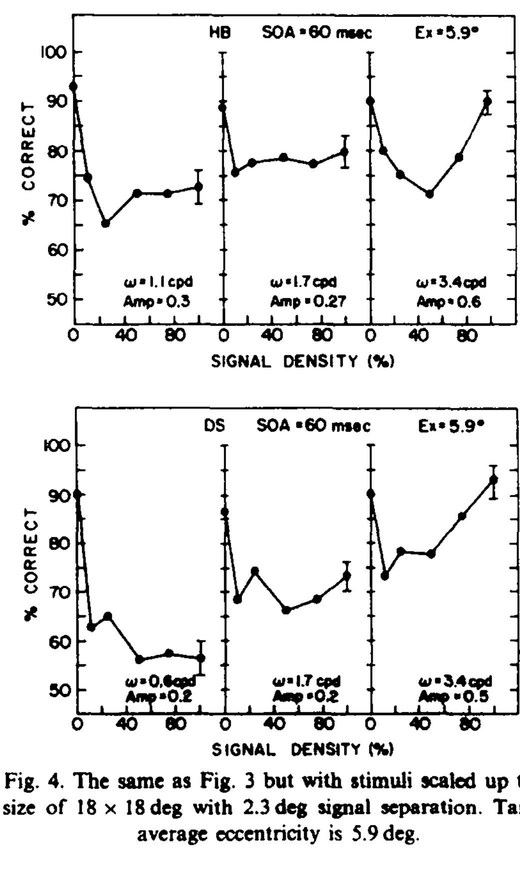

Fig. 4. The same as Fig. 3 but with stimuli scaled up to a size of 18 × 18 deg with 2.3 deg signal separation. Target average eccentricity is 5.9 deg.

separation (the larger separation is four times the Gaussian spread). In addition a constant gaussian spread implies increasing orientation bandwidth with decreasing spatial frequency. Thus, for the low spatial frequencies (< 2 c/deg) some overlap in the orientation spectrum is expected between the vertical and the horizontal patches (and both of them contain a d.c. component). The single target (density = 0) orientation resolution was kept constant (90% correct) for all frequencies, but if filter and/or hyper-filter integration area for the signals is larger than one or two cycles an interference between targets and distractors is expected, resulting a reduction in performance for high-density low spatial-frequency stimuli.

According to the data presented in Fig. 6, parallel processing efficiency is low for frequencies below 3/spacing c/deg (for all 8 curves), implying an integration radius of 3 cycles. This rule seems to hold for all stimuli whether sparse (3×3) or dense (9×9). The 3/spacing rule implies an integration area of 6 cycles which can be taken as an upper limit on integration (summation) area and thus to reflect a hyper-filter property. Filter integration area is probably smaller and may be about two cycles (Watson, Barlow & Robson, 1983). In a similar way, efficient parallel processing is limited to spatial frequencies below 9/spacing c/deg (this estimate is based on the large separation data from Fig. 6, and is consistent with all the data). The 9/spacing rule implies a total hyper-filter area of 18 cycles, including the inhibitory surround.

In summary, the data clearly show a non-monotonic dependence of detection performance on signal density. This is evident in the increased performance when signal density is high, an increase that is much stronger than observed before (Sagi & Julesz, 1987). In addition, it is demonstrated that this high density increase is sensitive to spatial separation in a frequency dependent way.

DISCUSSION

Experiments were described showing the effect of targets' spatial distribution on the detection of feature differences. It was demonstrated that increasing distractors' density can improve target detection dramatically. This improvement was found to be dependent on the spatial-frequency of the signals and their spatial separation. Best performance was achieved when the spatial separation between signals was

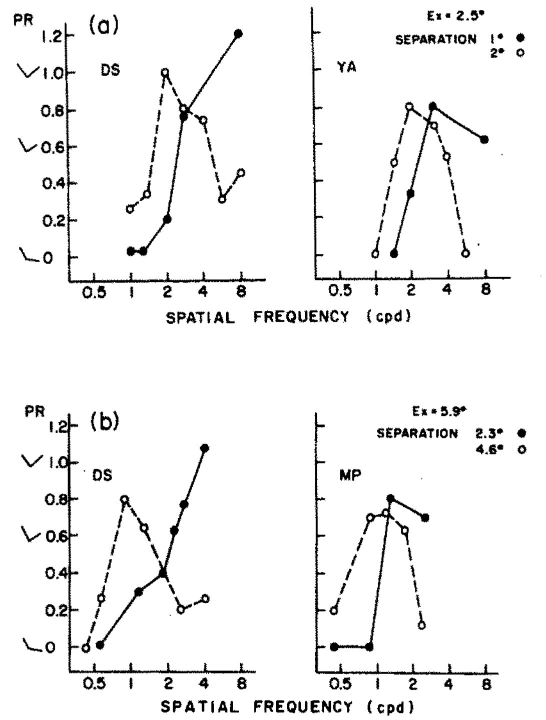

Fig. 6. The dependence of the parameter PR on spatial frequency for the two display sizes used 18×18 deg (a) and 8×8 deg (b). Closed symbols for dense displays (7×7) and empty symbols for sparse displays (3×3). Note that parallel detection efficiency (PR) increases with increasing spatial-frequency, however, for large patch separations it reaches a maximum at 2 c/deg (a) or at 1 c/deg (b) and then declines again.

larger than three times their center wavelength but smaller than nine times their wavelength.

The data support the idea that the mechanism underlying the detection of feature-differences operates through local interactions. The range of the interactions seems to scale with the wavelength of the filters operating in the detection task and has a radius of 9 times the filter typical wavelength. These interactions may have an antagonistic center-surround organization with a center radius of three times the filter typical wavelength. The range found here is larger than suggested from previous works (Sagi & Hochstein, 1985; Sagi & Julesz, 1987). This may be a result of the scaling effect found in the present study, the longer range interactions are due to the low spatial-frequency filters. However, it is still possible that these interactions have an absolute limit, thus a breakdown of the scaling effect may be expected below some spatial frequency. Since filter size depends on eccentricity (Wilson, 1983) the interaction range and limit may depend on eccentricity. Beck and Ambler (1973) found parallel performance on orientation (of the letter T) detection task for targets and distractors ($N = 8$) at an eccentricity of 18.2 deg with target size of 2.1 deg and separation of 13.8 deg. The size/spacing ratio (1:6.6) is well within the integration range found here (between 1:3 and 1:9), assuming that the T's are best detected by filters having wavelength corresponding to the T's size. (The assumption of low-frequency based detectability for alphanumeric kind of targets in brief displays seems to work in other cases as examined by Fogel & Sagi, 1989.) Previous studies that examined the role of element spacing used only line segments and thus there is no trivial scale for evaluation of optimal detection filter. However, since a filter wavelength larger than line length gives poor orientation sensitivity, line length can be taken as an upper limit on filter wavelength involved in the detection process. Line length was found to interact with spacing in determining texture discrimination performance (Nothdurft, 1985). The length/spacing ratio obtained by Nothdurft (1985) as an upper limit on texture discrimination were around 2:7, while Sagi and Julesz (1987) using a search task obtained an upper limit of 2:5. These ratios are on the lower end of the efficient range found here. It should be noted that the criteria for efficient parallel processing taken here is somewhat high, some of the low efficiency cases show a behavior that is usually taken as an evidence for parallel processing (Bergen & Julesz, 1983a). The term super-parallelism may be better suited for the phenomenon explored here.

The present results are consistent with a visual system organized hierarchically. In this hierarchy the first layer, composed of known spatial filters (spatial-frequency and orientation), is followed (after a nonlinearity) by a second layer that receives inhibitory and excitatory inputs from cells in the first layer according to a spatially antagonistic center-surround weighting function. A cell (hyper-filter) in the second layer should receive inputs from cells (filters) in the first layer having the *same properties*, except for spatial location (and phase). The number of filters feeding into the center of a hyper-filter may be estimated based on existing estimates of filter sizes. This estimate centers around two cycles for a filter (Caelli & Moraglia, 1985; Watson et al., 1983), thus approximately nine such filters would cover the central area of a hyper filter. Activity at the hyper-filter level accounts for the experimental results showing parallel detection of feature differences. A second-layer cell (hyper-filter) will not respond to a uniform texture within its receptive field, it will be active only if its excitatory and inhibitory inputs are not balanced, a case which will occur at perceivable texture boundaries. Texture boundaries created by more than one feature difference will be signaled by more than one type of hyper-filter; however, for automatic detection of texture boundaries it is sufficient to detect any activity at the hyper-field level. Note that texture uniformity should be defined with respect to the support of hyper-filters in the orientation/spatial-frequency/space domain. Thus, textures generated by replicating texture elements while randomizing their orientation are not uniform (in general) and may give rise to spurious texture boundaries. This scheme accounts very well for known search and texture segregation data (Fogel & Sagi, 1989; Rubenstein & Sagi, 1990).

Another alternative account for the present data is one that assumes a single layer network with mutual excitatory and inhibitory connections (Engle, 1974; Grossberg, 1987; Grossberg & Mingolla, 1985; Koch & Ullman, 1985; Sagi & Hochstein, 1985). According to our data most connections should be local, and the network may be viewed as a cooperative (over very short range) competitive (over somewhat larger range) system with spatial filters serving as basic units. The single layer scheme differs from the double layer hierarchical scheme by using feedback connections within the layer. The hierarchical scheme uses only feedforward connections. While there is no clear experimental evidence to support one account or the other, the two-layer scheme seems to be conceptually simpler.

The suggestion that detection of texture differences is performed by local measurements is consistent with one by Beck (Beck, 1972; Beck, Prazdny & Rosenfeld, 1983) and by Julesz (1986), however the present scheme avoids using structural elements (in addition to standard filters) and grouping operations. Also the detection of texture boundaries is not achieved here by measuring feature differences on a single dimension, like orientation difference, (Beck et al., 1983) or attribute based statistics (Voorhees & Poggio, 1988) but by measuring differential activity across similar filters. However, the scale over which this differential activity is measured as implied from the present study, seems to be close to the scale used by Voorhees and Poggio (1988) and Julesz (1986).

It is interesting to compare the experimental results obtained here with studies showing that the perceived contrast of a region in the visual field depends on the contrast in its surrounding neighborhood (Chubb, Sperling & Solomon, 1989; Sagi & Hochstein, 1985). The apparent contrast of a grating patch (having a given contrast) is inversely related to the contrast in its neighborhood. Chubb et al. (1989) found that this induction effect disappears when the surrounding and test regions have nonoverlapping spatial frequency spectra. They suggested that the response of each spatial filter is normalized relative to the responses of nearby filters of the same type, at a stage earlier to determination of apparent lightness (or contrast). Examination of the displays used here (Fig. 1) may suggest that if this normalization takes into account only responses of filters having similar orientation sensitivity, target and distractors will have different apparent contrast. The filter responding to the vertical Gabor signal will be normalized relative to its neighboring filters which are not stimulated. However, filters responding to horizontal Gabor signals will be normalized relative to a variable neighborhood activity, which depends on signal density. At low density of Gabor signals, or at large inter-signal distances, the different filters responding to the vertical and the horizontal signals will be normalized in a similar way and thus both targets and distractors will give rise to equally perceived contrast. Once distractor density is high, the normalization of their corresponding filters' response will tend to reduce the apparent contrast of the distractors, but not of the targets. If detection is based on apparent contrast only, and not on signal orientation, error rate is expected to decrease as the number of distractors increases (above some critical number). This scheme relies on the assumption that the target detection process cannot have parallel access to filters' labels, but only to their response value (see Introduction).

Finally, it is not clear what is the role of the suggested hyper-filters in object recognition (attentive vision). According to current views on attention feature-differences are detected as cues for the presence of objects (separating figure from ground), but then attentive vision takes over. However, it was shown that location of feature differences are available in parallel (Sagi & Julesz, 1985) and some shape information is available from this information (observers could discriminate between triangles defined by three feature-gradients at the time these gradients were detected). Braun and Sagi (1990) showed that information about feature-gradients is available without using attention, thus raising the possibility that feature-gradients may provide a base representation for a recognition system that operates concurrently with attentive vision. Since hyper-filters provide rich information concerning feature and luminance gradients they may serve as an input stream to a fast mechanism that can identify global properties of objects without the temporal limitations imposed on attentive vision.

Acknowledgement—I would like to thank Jochen Braun and Barton Rubenstein for their helpful comments. This work was supported by the Basic Research Foundation administered by the Israel Academy of Sciences and Humanities.

REFERENCES

Arguin, M. & Cavanagh, P. (1988). Parallel processing of two disjunctive targets. *Perception and Psychophysics, 44(1)*, 22–30.

Barlow, H. B. (1981). Critical limiting factors in the design of the eye and the visual cortex. *Proceedings of the Royal Society of London B, 212*, 1–34.

Beck, J. (1972). Similarity grouping and peripheral discriminability under uncertainty. *American Journal of Psychology, 85*, 1–19.

Beck, J. & Ambler, J. (1972). Discriminability of differences in line slope and in line arrangement as a function of mask delay. *Perception and Psychophysics, 12*, 33–38.

Beck, J., Prazdny, K. & Rosenfeld, A. (1983). A theory of textural segmentation. In Beck, J. & Rosenfeld, A. (Eds.), *Human and machine vision*. New York: Academic Press.

Beck, J., Sutter, A. & Ivry, R. (1987). Spatial frequency channels and perceptual grouping in texture segregation. *Computer Vision, Graphics and Image Processing, 37*, 299–325.

Bergen, J. R. & Adelson, E. H. (1988). Early vision and texture perception. *Nature, London, 333*, 363–364.

Bergen, J. R. & Julesz, B. (1983a). Parallel versus serial processing in rapid pattern discrimination. *Nature, London, 303*, 696–698.

Bergen, J. R. & Julesz, B. (1983b). Rapid discrimination of visual patterns. *IEEE, SMC-13 (5)*, 857–863.

Braun, J. & Sagi, D. (1990). Vision outside the focus of attention. *Perception and Psychophsics*, in press.

Caelli, T. M. & Moraglia, G. (1985). On the detection of Gabor signals and discrimination of Gabor textures. *Vision Research, 25*, 671–684.

Campbell, F. W. & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, London, 197*, 551–566.

Chubb, C., Sperling, G. & Solomon, J. (1989). Texture interactions determine apparent lightness. *Proceedings of the National Academy of Sciences, U.S.A.*, *86*, 9631–9635.

Daugman, J. (1984). Spatial visual channels in the Fourier plane. *Vision Research, 24*, 891–910.

Daugman, J. D. (1985). Uncertainty relation for resolution in space, spatial frequency and orientation, optimized by two dimensional cortical filters. *Journal of Optical Society of America A*, *2*, 1160–1169.

Daugman, J. D. (1987). Image analysis and compact coding by oriented 2D Gabor primitives. *SPIE Processing*, *758*, 19–30.

Dick, M., Ullman, S. & Sagi, D. (1987). Parallel and serial processes in motion detection. *Science*, *237*, 400–402.

Duncan, J. (1985). Visual search and visual attention. In Posner, M. I. & Mann, O. S. M. (Eds.), *Attention and performance XI* (pp. 85–106). Hillsdale, New Jersey: Erlbaum.

Engle, F. L. (1974). Visual conspicuity and selective background interference in eccentric vision. *Vision Research*, *14*, 459–471.

Fogel, I. & Sagi, D. (1989). Gabor filters as texture discriminator. *Biological Cybernetics*, *61*, 103–113.

Graham, N. (1989). Low-level visual processes and texture segregation. *Physica Scipta*, *39*, 147–152.

Graham, N. & Nachmias, J. (1971). Detection of grating pattern containing two spatial frequencies: A comparison of single-channel and multi-channel models. *Vision Research*, *11*, 251–259.

Graham, N. & Robson, J. G. (1987). Summation of very close spatial frequencies: The importance of spatial probability summation. *Vision Research*, *27*, 1997–2007.

Grossberg, S. (1987). Cortical dynamics of 3D form, color, and brightness perception: 1. Monocular theory. *Perception and Psychophysics*, *41(2)*, 87–116.

Grossberg, S. & Mingolla, E. (1985). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception and Psychophysics*, *38(2)*, 141–171.

Gurnsey, R. & Browse, R. A. (1987). Micropattern properties and presentation condition influencing visual texture discrimination. *Perception and Psychophysics*, *41(3)*, 239–252.

Julesz, B. (1981). Textons, the elements of texture perception and their interactions. *Nature, London*, *290*, 91–97.

Julesz, B. (1986). Texton gradients: The texton theory revisited. *Biological Cybernetics*, *54*, 464–469.

Julesz, B. & Kröse, B. (1988). Features and spatial filters. *Nature, London*, *333*, 302–303.

Koch, C. & Ullman, S. (1985). Shifts in visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, *4*, 219–227.

Kröse, B. J. A. (1987) Local structure analyzers as determinants of preattentive pattern discrimination. *Biological Cybernetics*, *55*, 289–298.

Landy, M. S. & Bergen, J. R. (1989). Texture segregation for filtered noise patterns. *Investigative Ophthalmology and Visual Science* (Suppl.), *30*, 160.

Legge, G. E. (1979). Spatial frequency masking in human vision: Binocular interactions. *Journal of the Optical Society of America*, *69*, 838–847.

Malik, J. & Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America A*, in press.

Nakayama, K. & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, London*, *320*, 264–265.

Nothdurft, H. C. (1985). Sensitivity for structure gradients in texture discrimination tasks. *Vision Research*, *25*, 1957–1968.

Nothdurft, H. C. (1988). Masking of textons. *European Brain and Behavior Society Workshop*, Tubingen, March 1988.

Nothdurft, H. C. (1990). Texton segregation by associated differences in global and local luminance distribution. *Proceedings of the Royal Society of London B*, *239*, 295–320.

Pashler, H. (1988) Cross-demensional interaction and texture segregation. *Perception and Psychophysics*, *43(4)*, 307–318.

Rentschler, I. & Treutwein, B. (1985). Loss of spatial phase relationships in extrafoveal vision. *Nature, London*, *313*, 308–310.

Rubenstein, B. S. & Sagi, D. (1990). Spatial variability as a limiting factor in texture discrimination tasks: Implications for performance asymmetries. *Journal of the Optical Society of America A*, in press.

Sagi, D. (1988). The combination of spatial frequency and orientation is effortlessly perceived. *Perception and Psychophysics*, *43*, 601–603.

Sagi, D. & Hochstein, S. (1983). Discriminability of suprathreshold compound spatial frequency gratings. *Vision Research*, *23*, 1595–1606.

Sagi, D. & Hochstein, S. (1985). Lateral inhibition between spatially adjacent spatial frequency channels? *Perception and Psychophysics*, *37(4)*, 315–322.

Sagi, D. & Julesz, B. (1985). "Where" and "What" in vision. *Science, New York*, *228*, 1217–1219.

Sagi, D. & Julesz, B. (1987). Short range limitations on detection of feature differences. *Spatial Vision*, *2*, 39–49.

Stromeyer, C. F. & Julesz B. (1972) Spatial-frequency masking in vision: Critical bands and spread of masking. *Journal of the Optical Society of America*, *62*, 1221–1232.

Treisman, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics and Image Processing*, *31*, 156–177.

Treisman, A. & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97–136.

Turner, M. R. (1986). Texture discrimination by Gabor functions. *Biological Cybernetics*, *55*, 77–82.

Voorhees, H. & Poggio, T. (1988). Computing texture boundaries from images. *Nature, London*, *333*, 364–367.

Walters, D., Biederman, I. & Weisstein, N. (1983). The combination of spatial-frequency and orientation is not effortlessly perceived. *Investigative Ophthalmology and Visual Science* (Suppl.), *24*, 238.

Watson, A. B. (1983). Detection and recognition of simple spatial forms. In Braddick, O. J. & Sleigh, A. C. (Eds.), *Physical and biological processing of images*. Berlin: Springer.

Watson, A. B. & Robson, J. G. (1981). Discrimination at threshold: Labeled detectors in human vision. *Vision Research*, *21*, 1115–1122.

Watson, A. B., Barlow, H. B. & Robson, J. G. (1983). What does the eye see best? *Nature, London*, *302*, 419–422.

Wertheim, A. H. (1981). Distraction in visual search. Report IZF 1981-7, Institute for perception TNO, Soesterberg (The Netherlands).

Wilson, H. R. (1983). Psychophysical evidence for spatial channels. In Braddick, O. J. & Sleigh, A. C. (Eds.), *Physical and biological processing of images*. Berlin: Springer.

Wilson, H. R. & Bergen, J. R. (1979). A four mechanism model for threshold spatial vision. *Vision Research*, *19*, 19–32.