

Spatial variability as a limiting factor in texture-discrimination tasks: implications for performance asymmetries

Barton S. Rubenstein and Dov Sagi

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

Received August 21, 1989; accepted February 2, 1990

Texture-discrimination tasks reveal a pronounced performance asymmetry depending on which texture represents the foreground region (small area) and which represents the ground (large area). This asymmetry implies that some global processes are involved in the segmentation process. We examined this problem within the context of the texture-segmentation algorithm, assuming two filtering stages. The first stage uses spatial frequency and orientation-selective (Gabor) filters, whereas the second stage is formed by low-resolution edge-detection filters. The presence and location of texture borders are indicated by significant responses in the second stage. Spurious texture borders may occur owing to textural local variabilities (such as orientation randomization), which are enhanced by the first stage. We suggest that these spurious borders act as background noise and thus limit performance in texture-discrimination tasks. The noise level depends on which texture occupies the ground in the display. We tested this model on numerous pairs of textures and found remarkably good correlation with human performance. A prediction of the model, namely, that discrimination asymmetry will be reduced when textural elements have identical orientation, was tested psychophysically and confirmed.

INTRODUCTION

Texture-discrimination tasks reveal a pronounced performance asymmetry in respect to which texture represents the foreground (disparate) region and which represents the ground (large area).¹ In other words, two different textural elements that can be discriminated occasionally have a unique characteristic that makes one texture more salient than the other (see Fig. 1). This characteristic implies that more than local or adjacent textural element analysis occurs in texture-discrimination tasks.²⁻⁵ Otherwise, asymmetric discrimination would not exist. Rather, some global processes must be involved in the discrimination and segmentation process.

We propose a solution to the problem that is based on simple statistical information that is inherent in all textural stimuli. We claim that these textures can be described as having distinct noise characteristics that indicate how easily a texture region can be detected when embedded in another texture. Thus, if simple signal-in-noise arguments are applied, a texture that is surrounded by a noisy background (composed of another texture) will be harder to detect than with a less noisy background. If, for example, a Γ shape (Fig. 1) has more noise characteristics than a $+$ shape, then we would expect it to be more difficult to detect the $+$'s among the noisier Γ 's in the background than in the reverse case.

This proposal would also imply that textures having similar noise characteristics would have symmetric performance levels. Finally, the most important claim is that noise characteristics are caused primarily by spatial variability properties of the stimulus. Hence spatial variability acts as a direct limiting factor in the discrimination process, or, in other words, asymmetric performance is controlled by these noise characteristics. We shall analyze

these characteristics, but, before doing this, we review various different types of asymmetry so as to gain a better perspective of the problems of texture discrimination.

Segmentation asymmetry has been observed in many cases^{1,6-8}; however, only a few cases have been classified according to stimulus characteristics. These characteristics include differences in element intensities, positional variability, and a special case treated by Treisman^{6,7} that involves vertical and near-vertical lines. In a more fundamental discrimination task, two elements regardless of their compositional and positional characteristics can be compared simply by their light intensities. Elements that are brighter will stand out much more among relatively dark elements than in the reverse case. This difference in our ability to see is called asymmetry. Weber's law predicts asymmetry of such a discrimination task, pointing to the brighter elements as the more salient features. Weber's law states that detectability of a target will depend on the adaptation level produced by the stimulus (background). When a dark background is presented, a small intensity change will be easily detected, whereas a bright background will necessitate a larger detection threshold. Hence this phenomenon may be a result of different visibilities of the two textures. Other attempts to model this phenomenon include filtering algorithms (difference-of-Gaussian and Gabor) that are sensitive to contrast and hence make intensity-based discrimination a simple task.

Asymmetry can also be caused by elements with different positional characteristics (jitter). Gurnsey and Browse describe similar elements with different jitter components and point to the noisier ones (more jitter) as the more salient. Again with filtering models, difference-of-Gaussian and Gabor filters are sensitive to changes in the texture power spectrum introduced by jitter, rendering discrimination a

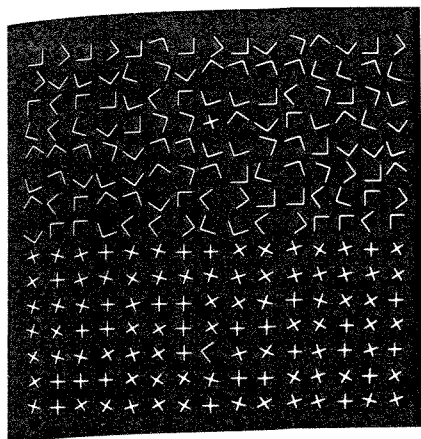


Fig. 1. Example of asymmetry. Visual examination shows 'L' 's in the foreground (bottom half) as more salient.

straightforward task. Our explanation is that jitter components can be viewed as noise characteristics, such that a texture embedded in a noisy background (considerable jitter) would be more difficult to detect than a background having less spatial jitter.

Treisman and co-workers^{6,7,9,10} also speak about an asymmetry in the case of vertical and near-vertical lines, which is quite different from the cases mentioned above. She claims that near-vertical lines are more salient than vertical ones (with each line having the same length). Treisman's explanation of this asymmetry is an intuitive one, based on the notion that nonvertical lines are special features in the visual field compared with vertical lines, which are so common in our world that they are relatively unnoticed. Hence, because features stand out, the nonvertical line will be the more salient of the pair. Alternatively, we speculate that filters responding to tilted lines are noisier than those responding to vertical lines. Hence tilted lines in the background will generate noisier backgrounds, and, as a result, detectability of vertical targets among tilted distractors will be reduced. As will be shown below, this explanation is in accordance with the structure of our proposed model.

Finally, the most subtle discrimination task includes elements having identical intensities and similar positional variability (jitter). To explain this, Julesz^{11,12} suggests that elongated blobs having particular orientation, lengths, widths and intensities, line terminations, and crossings are separate textons, which act as fundamental elements in textural discrimination and in the segmentation process. Texton theory states that discrimination is possible only if the number of textons is different. For example, both H and L have three segments, but the former has four terminators compared with two for the latter; hence discrimination is possible. Further, Julesz^{11,12} states that the element with the most textons will be more salient (H in this case). This theory works fairly well in modeling this phenomenon; however, there are some inconsistencies that point to limitations of such a model. Contrary to this structural approach, filtering algorithms have until now been unsuccessful in predicting discrimination asymmetry. In the present paper, however, we propose another filtering model that is based on a statistical and biological intuition, using special characteristics of the orientation-selective linear filters to model this peculiar phenomenon. This approach, unlike that of other

models, is capable of modeling all the above-mentioned asymmetries, with the addition of some other examples, shown below.

2. PROPOSED MODEL

Our model is based on one described by Fogel and Sagi¹³ that can predict psychophysical performance levels for discrimination tasks involving textures. This texture-discrimination algorithm produces a set of filtered images (maps), with each map representing the spatial distribution of image energy contained within a specific spatial frequency and orientation band. Gabor filters, weighting functions that are tuned to spatial location, orientation, and spatial frequency, are the fundamental tools of this task and are used to discriminate between foreground and background. In general, filtered foreground and background textures produce maps that are discriminated quite satisfactorily, in the sense that activity differs, on average, between foreground and background whenever the textures are psychophysically discriminable.¹³⁻¹⁷ However, we find that activity levels within either foreground or background are at times not uniform but show considerable fluctuation. In this situation, the reliability of texture segmentation depends on the magnitude of the foreground-background difference of filter energies (signal) relative to the amplitude of fluctuations in the background (noise). If filter energy fluctuation amplitudes differ between foreground and background textures, signal-detection theory predicts a segmentation asymmetry.

A number of factors may contribute to the spatial nonuniformity of a texture and the resulting activity fluctuations. A major source of spatial nonuniformity is the randomization of the orientation of textural elements. Random element orientation will have the result that the orientation of neighboring elements is more similar in some parts of the texture than in others. Hence this orientation variability is being transformed into spatial variability. Since different textural elements have different orientation variability, they must also have different spatial variability, and, because the latter variability limits performance (by acting as background noise), asymmetry should occur. It should be pointed out that orientation randomization was introduced to the discrimination task in order to permit second-order statistics of foreground and background textures to be equated^{18,19} and thus is justified within the statistical framework used in the earlier studies of Julesz.

Gabor filters are, by design, orientation and spatial-frequency selective and spatially constrained and therefore can be said to mimic similar characteristics of cortical simple cells.²⁰ This is not to say that neurons act as Gabor filters but only that certain low-level processes in the brain are modeled well by Gabor filters. Psychophysical studies strongly support the existence of orientation- and spatial-frequency-selective filters in early vision.²¹⁻²⁵ Furthermore, it was shown that the visual system can detect in parallel only the location of feature gradients (orientation changes of adjacent texture elements as an example), whereas the identification of gradients requires serial search.^{4,5} This implies that spurious local changes existing within the background texture have to be considered in search and texture-segmentation tasks.²³ On the neurophysiological level, cells were found in areas V1 and V2 of the macaque monkey that

respond best to local orientation changes and not to uniform textures.²⁶ These findings imply the existence of a second stage filtering in early vision, where the initially filtered (Gabor or similar) image is filtered again in order to find local energy differences.²⁷⁻³⁰

The segmentation asymmetry predicted by the present model is based on the variability in the response of Gabor filters to textural elements at different orientations, thus generating local energy differences. Certain elements, for example, + and Γ , have different energy profiles when filtered by Gabor filters at different orientations. This indicates not only that these elements have different energy profiles across the spatial frequency spectrum as already suggested^{13,31-33} but that each element has a unique variability profile across the orientation spectrum. We show here that the variability profile of a textural element across the orientation spectrum predicts its salience, and this correlates well with the psychophysical experiments (concerning asymmetry) of Gurnsey and Browse.¹ This result indicates the importance of spatial variability in discrimination tasks.

To justify these claims, we prepared a computer simulation of the discrimination task, so as to study visually the importance of the orientation variability and hence of the spatial variability. Then, based on the success of these results, we developed a mathematical model that is able to predict human performance. This model is a refinement of the discrimination-segmentation model proposed by Fogel and Sagi.¹³ It incorporates spatial and orientation variability as variables in the discrimination process, using the algorithm from the computer simulation as its fundamental building blocks. Finally, to test our hypothesis directly we performed psychophysical experiments, controlling orientation variability within textures.

3. ANALYSIS OF SPATIAL VARIABILITY

A. Variability across the Orientation Spectrum

First we show some concrete correlation between variability of the orientation profiles of elements and their relative salience according to psychophysical results of Gurnsey and Browse.¹ Using even and odd Gabor filters having fixed orientation and spatial frequency, we computed the filtered energy of various elements. The equations for these filters are

$$G_{ev}(x, y|\lambda, \theta, x_c, y_c) = \exp\left\{\frac{-[(x - x_c)^2 + (y - y_c)^2]}{2\sigma^2}\right\} \times \cos\{(2\pi/\lambda)[(x - x_c)\cos\theta - (y - y_c)\sin\theta]\}, \quad (1)$$

$$G_{od}(x, y|\lambda, \theta, x_c, y_c) = \exp\left\{\frac{-[(x - x_c)^2 + (y - y_c)^2]}{2\sigma^2}\right\} \times \sin\{(2\pi/\lambda)[(x - x_c)\cos\theta - (y - y_c)\sin\theta]\}, \quad (2)$$

where σ is the Gaussian width ($\sigma = 8$), θ is the filter orientation and λ is its wavelength, and x_c and y_c represent the center of the filter with x and y being the coordinates over the filter's domain. Let $L(x, y)$ be the input matrix representing one element and its surrounding space (32×32 pixels) and Eqs. (1) and (2), the Gabor operators. Then, by

performing a simple dot product, $G \cdot L$, we produce two response values:

$$GL_{ev}(x_c, y_c|\lambda, \theta) = \sum_{x, y} G_{ev}(x, y|\lambda, \theta, x_c, y_c) \cdot L(x, y),$$

$$GL_{od}(x_c, y_c|\lambda, \theta) = \sum_{x, y} G_{od}(x, y|\lambda, \theta, x_c, y_c) \cdot L(x, y),$$

where x and y are indices over the basic matrix element = 0-31). Finally, a simple squaring and summing of convolutions of Eqs. (3) and (4) gives the combined texture energy level:

$$E(x_c, y_c|\lambda, \theta) = GL_{ev}^2(x_c, y_c|\lambda, \theta) + GL_{od}^2(x_c, y_c|\lambda, \theta).$$

Since Eq. (5) produces only a locally shift-invariant response, we add a spatial integration stage (over x_c, y_c) obtain a texture energy measure:

$$T_{x, y}(\lambda, \theta) = \sum_{x_c, y_c \in S(x, y)} E(x_c, y_c|\lambda, \theta),$$

where x_c and y_c are indices indicating the center of the filter and $S(x, y)$ is some excitatory neighborhood corresponding to a smoothing operation. This neighborhood is defined to ensure spatial invariance, and hence energy levels for different filters are summed, with filter centers separated from one another equally in a 5×5 matrix. This energy represents the shift-invariant response of the system. Figure 2 is a plot of energy T for two different elements, + and Γ . (Each element is composed of line segments 17 pixels long.) Each graph shows energies (plotted vertically) for different orientations and filter wavelengths. Notice that for filters with large wavelengths (backs of curves) the + has little variability, while the Γ has much more. Correspondingly, because the Γ is noisier than the +, +s should be more salient in the disparate region than in the reverse case, as shown by Gurnsey and Browse.¹ Also note that, for short wavelengths, there is great sensitivity to individual line elements; but we have found that energy curves of wavelengths larger than element size ($\lambda > 17$) correlate more closely with psychophysical data. In fact, as will be shown explicitly in Section 4, nearly all pairs of elements presented by Gurnsey and Browse¹ expose asymmetries that correlate closely with variability characteristics across the orientation spectrum. Elements that were more salient psychophysically also showed the most orientation variability (in the large-wavelength window) with respect to the complement element of the pair.

B. Computer Simulation

Based on these preliminary studies, we created a three-stage algorithm that simulates textural segmentation. The stimulus consists of randomly oriented +s and Γ 's [Figs. 3(a) and 4(a)]. The first step of the algorithm involves filtering the image with a Gabor filter of a specific orientation and spatial frequency ($\theta = 72^\circ$, $\lambda = 19$ pixels, $\sigma = 8$ pixels). Specifically, we chose a 32×32 matrix filter and applied it to corresponding 32×32 pixel patches of the stimulus at a 4-pixel sampling rate. With this small sampling rate there was a significant overlap of the filtering, thus removing any phase properties of filter placement on the image. [Note that each element of Figs. 3(a) and 4(a) has a

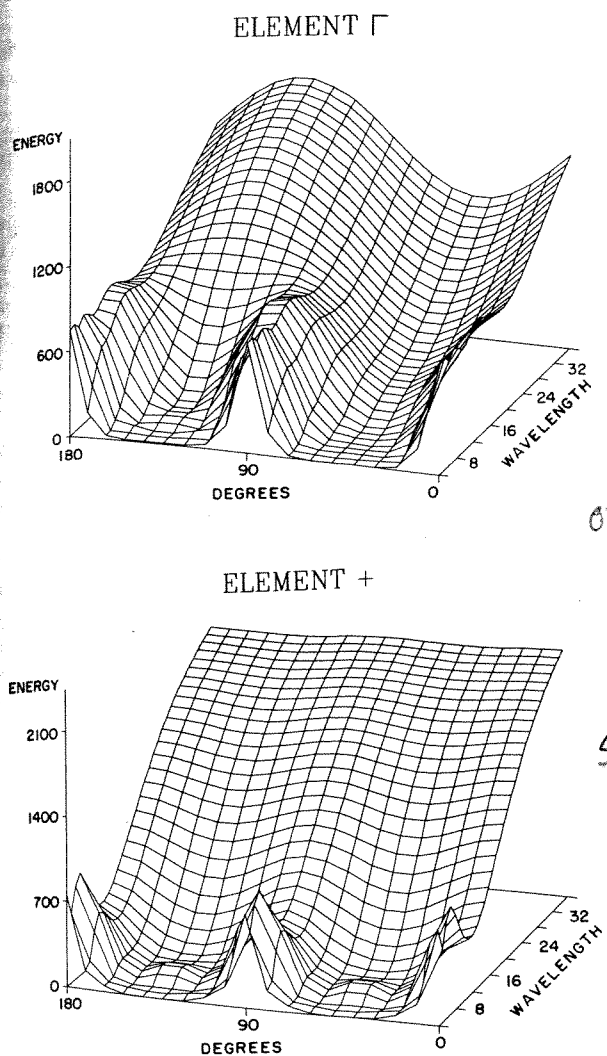


Fig. 2. Gabor filter texture energies, $T(\lambda, \theta)$, of + and Γ for different wavelengths and orientations. Notice that, for large wavelengths, element Γ has greater energy variability across the orientation spectrum.

segment size of 17, sitting within a 32×32 matrix of space.] Figures 3(b) and 4(b) represent the output of this energy filtering. Notice the increased variability of energies of the Γ 's relative to those of the + 's. By looking at the Gabor-filtered images, one can see how the figure representing Γ 's in the foreground [Fig. 4(b)] is easier to segment.

The second stage of the algorithm involves filtering the Gabor energy image with a two-dimensional Gaussian filter. The technique is similar to the first stage, with a 128×128 matrix filter being applied at a 4-pixel sampling rate. Here, $\sigma = 23$ pixels. The application of such a large filter produces a smoothed image with little local variability of energy levels, except, of course, where large local differences exist. Figures 3(c) and 4(c) represent these filtered outputs. Finally, for the third stage, significant local intensity gradients are detected from the Gaussian-filtered images, producing Figs. 3(d) and 4(d). This stage detects local differences in energy levels and hence is a type of edge-detection algorithm. Specifically, this was done by thresholding the Gaussian maps [Figs. 3(c) and 4(c)] and then running an

edge-detection algorithm on the resultant two-value map. Notice that in Fig. 4(d) edge detection is clear and pronounced (Γ 's in foreground). In contrast, Fig. 3(d) shows detection of the stimulus with additional spurious edges in the background. These spurious edges represent the energy variability across the orientation spectrum of the Γ 's. Specifically, Γ 's with similar orientations tend to group together because adjacent Γ 's of different orientations have large differences in activity. Likewise, because + 's have little variability, groups of + 's distinguishing themselves from other + 's are rare, as can be seen in the background of Fig. 4(d). We conclude that these noise characteristics that pro-

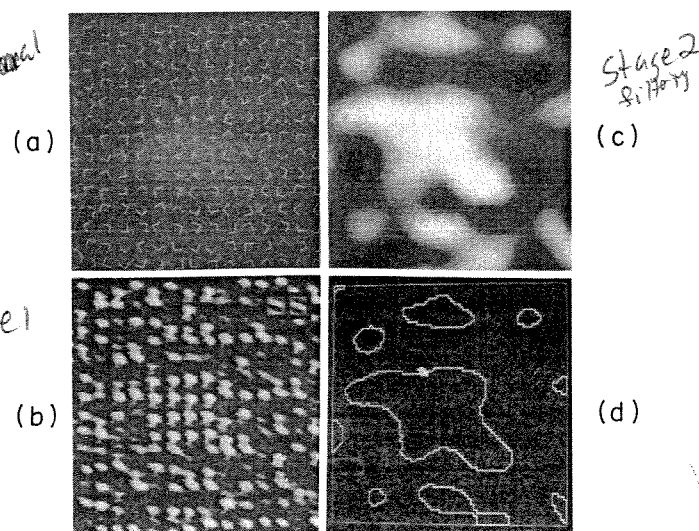


Fig. 3. Computer simulation of texture segmentation. The stimulus (a) is acted on by even and odd Gabor filters having an orientation of 72° to produce an energy map (Gabor patch also shown) (b). This image is in turn smoothed by a Gaussian filter (c), and finally the Gaussian image is thresholded and edge detection is performed (d). Notice the spurious clusters caused by the Γ 's in the background.

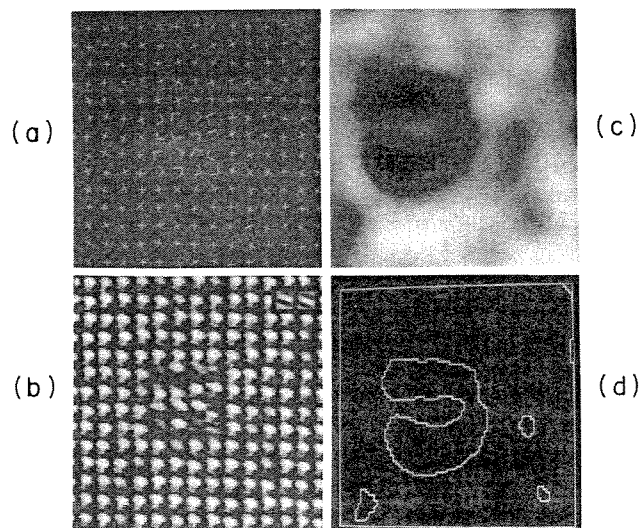


Fig. 4. Same as Fig. 3 except with Γ 's in the foreground. Notice the relatively less noisy background caused by the + 's having limited orientational variability.

tion of the specific element energy at a specific frequency. Two different distributions should be considered, one for the foreground and one for the background. For the sake of simplicity, we assume that these two distributions are normal, with means M_f and M_b and variances V_f and V_b for foreground and background, respectively, where M_f , M_b , V_f and V_b can be computed from the orientation distribution at a specific frequency. (Although the assumption of normality is not correct, it may be justified by considering a broad spatial averaging process that smoothes the filtered image. Hence, assuming textures with randomly oriented elements, summing log-energy distributions belonging to adjacent elements will produce an approximately normal distribution.)

Orientation variability is not the only noise source in the system. We have to consider internal noise (filter response variability) and noise introduced by the mask used in the experiments. Practically, the mask used limits performance according to its temporal delay from the stimulus; thus the temporal dimension must be considered. However, since we do not consider the temporal dimension here, we assume that all noise sources can be described by one normal distribution with a mean of M_n and a variance of V_n . M_n depends on the mask and the filter's spatial and temporal properties; it is probably frequency dependent according to the different temporal properties of the different filters.³⁶ However, M_n should not depend on the particular textures used, and since we will consider only filter response differences, the mean noise level cancels out. We assume that V_n , which suffers from the same problems as M_n , is the same for all filters (of different spatial frequency and orientation) and independent of the texture elements used. Note that, in order to account for Treisman's result concerning vertical versus tilted line detection,^{6,7} we should make V_n smaller for vertically oriented filters. Since we do not have any straightforward way of estimating the noise variance V_n , we take it as the only free parameter for adjusting the model performance in relation to human performance.

Both the noise distribution and the orientation distribution produce spatial variability. If we assume that the two distributions are normal, and uncorrelated, we get normal distributions of the sums with means $M_{fn} = M_f + M_n$ and $M_{bn} = M_b + M_n$ and variances $V_{fn} = V_f + V_n$ and $V_{bn} = V_b + V_n$ for foreground and background, respectively. (M_f and M_b represent the mean log energies described above.)

4. Computing Spatial Distributions of Local Differences

Here we obtain the distributions for F-B and B-B differences from the two distributions of the texture log-energy values. We assume that local differences are computed as differences between pairs of log-energy values, $T_{x,y}$, separated by a distance larger than the size of the smoothing neighborhood, $S(x, y)$. Hence the B-B difference distribution p_{bb} has a zero mean and a variance of

$$V_{bb} = 2V_{bn} = 2(V_b + V_n), \quad (9)$$

where V_{bb} is the variance of local differences samples in the background due to orientation randomization. The F-B difference distribution (at the F-B border) has a mean of

$$M_{fb} = M_{fn} - M_{bn} = M_f - M_b, \quad (10)$$

which reflects the difference between filter responses to the

foreground and background elements, averaged across all orientations. The variance of p_{fb} is

$$V_{fb} = V_{fn} + V_{bn} = V_f + V_b + 2V_n. \quad (11)$$

Since these two difference distributions are normal, they contain negative difference values. Thus, when computing percent-correct response [Eq. (6)] we have to transfer the negative values into the positive range (see Subsection 4.A).

Finally, the F-B and B-B absolute-value distributions are

$$p_{bb}(z) = \frac{2}{\sqrt{2\pi V_{bb}}} \exp\left(-\frac{z^2}{2V_{bb}}\right) \quad \text{for } z \geq 0 \quad (12)$$

and

$$p_{fb}(z) = \frac{1}{\sqrt{2\pi V_{fb}}} \left\{ \exp\left[-\frac{(z - M_{fb})^2}{2V_{fb}}\right] + \exp\left[-\frac{(-z - M_{fb})^2}{2V_{fb}}\right] \right\} \quad \text{for } z \geq 0, \quad (13)$$

where z denotes local-energy differences.

5. Integrating Information from Filters of Different Spatial Frequencies

The log-energy spatial distributions obtained above vary according to the filter's spatial frequency. Thus substituting Eqs. (12) and (13) into Eq. (7) yields a correct response value only when a specific spatial-frequency band is used. Integrating information from different filters is a nontrivial problem. We applied two (alternative) simple combination rules: (1) use the filter that yields the highest percent-correct value and ignore the others; (2) use only low-spatial-frequency filters. Application of the second rule does not require knowledge of the stimulus parameters. Application of the first rule requires isolation of the relevant filter, which may depend on the specific input.

C. Model Performance

Model performance was tested numerically. Energy values were calculated by using Gabor filters with $\sigma = 24$ pixels, an amplitude of 1, and a λ range from 8 to 64 pixels (at steps of 4 pixels). Convolutions were computed by summing 25 different Gabor energies with filter centers separated from one another by 16 pixels and positioned in a 5×5 matrix [defining the $S(x, y)$ neighborhood in Eq. (8)]. It was assumed that interelement structure does not play an important role in this computation (for sparsely spaced elements); thus convolution values were obtained for single elements. This assumption may be justified by the result of our psychophysical experiments (described in Section 5 below), with a spacing-to-element size ratio of 2.5 (Gurnsey and Browse³⁷ used a ratio of 2.3), showing a negligible effect of interelement structure. (This of course is not true for denser textures.) Performance was examined in each case (with each element of the pair in the foreground) across all spatial frequencies, and the highest performance percentage was selected as the model prediction. This maximum-over- λ strategy has little effect on the data, since in most cases the best performance was obtained by using low-frequency filters with wavelengths near $1.5(\pm 0.5)$ times element size. Element size was 33×33 pixels for square-shaped elements and 25×49 pixels

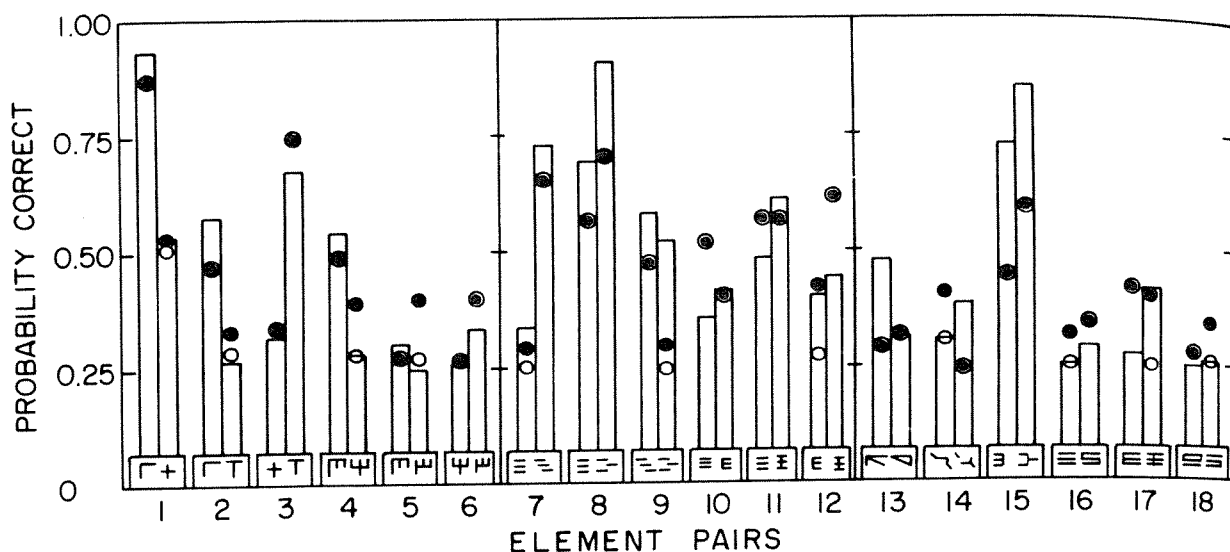


Fig. 5. Correlation between the psychophysical results of Gurnsey and Browse¹ (histogram) and the predictions of the model (circles). The figure is presented in groups of two elements (numbered at the bottom of each pair), representing a particular stimulus. Each histogram rectangle represents the psychophysical performance level with the element depicted below it in the foreground and the adjacent element in the background. Filled circles represent the prediction obtained by selecting the spatial frequency yielding the highest performance for each pair. The open circles represent predictions for Gabor filters having only larger wavelengths. Note that perfect correlation would place one class circle on top of each histogram rectangle. ($V_n = 0.05$.)

otherwise. Line segments were 3 pixels wide, similar to what was used in the psychophysical experiments.³⁷

As was noted above, the only free parameter is the noise variance V_n ; all other variables were calculated by applying the texture elements used. The model predictions are depicted in Fig. 5 ($V_n = 0.05$) along with the psychophysical data (histogram) of Gurnsey and Browse.¹ The figure is presented in groups of two elements (numbered at the bottom of the figure), representing a particular stimulus. Each histogram rectangle represents the psychophysical performance level with the element depicted below in the foreground and the adjacent element in the background. Likewise for the other histogram rectangle of the pair. The circles represent the corresponding predictions of our model. The filled circles represent the highest performance predicted for one spatial frequency from a range of frequencies $4 < \lambda < 64$. The open circles represent predictions for Gabor filters having only larger wavelengths $\lambda > 40$ pixels (1.25 times element size). Note that perfect correlation would place one class of circles on top of each histogram rectangle.

The model was also tested with other values of σ (16 and 32) and produced similar results with one exception: improvement of performance for pairs that differ in total pixel number (pair number 10 and 11) for $\sigma = 16$. At this σ value the most discriminating filters are low-pass filters and are thus more intensity sensitive.

A visual inspection of the data shows a surprisingly good correlation between the model performance and human performance. The percent-correct values produced by the model are close to the ones obtained in psychophysical experiments, and asymmetries, whenever they are pronounced, correlate remarkably well. The total correlation value is 0.80 ($r = 0.83$ for the low-frequency version). It should be noted that the psychophysical data points are based on a limited number of trials ($n = 384$) distributed

across twelve observers (which are not the same for all texture pairs) and across four stimulus durations; thus a perfect correlation when only one adjustable parameter is used is not expected.

D. Model Summary

We have presented a model of human texture discrimination that seems to account well for psychophysical results. Texture differences are detected here by differential activity orientation and spatial frequency filters. A decision rule was added in order to allow real texture borders to be separated from spurious ones that are caused by texture spatial variability and noise. This decision rule produces asymmetry in detection because of varying noise levels in the background caused by different elements.

We believe that this decision stage cannot be avoided, and thus asymmetry is expected from any model that is based on detection of local differences (without identification). We chose a simple detection rule, detection of a global maximum of local differences.

This hypothesis seems to work in our case, but more complicated decision rules are possible, depending on the complexity of the available information in the local-difference maps. In addition, the level of familiarity of the observers with the textured patterns must be considered, with the uncertainties introduced by the experimenter taken into account. In the experiments that we modeled, the different texture pairs were intermixed in blocks of trials; thus a general decision strategy had to be used. In another set of experiments, Gurnsey and Browse¹ trained two observers to improve their performance. Within the framework introduced here, this improvement can be a result of improved decision strategy, which can be accomplished by considering the size of the foreground as a detection facilitator, applying minima detection in addition to or instead of maxima detection.

tion, or looking for specific types of edge value. Fogel and Sagi¹³ overcame some of these problems by using an algorithm that removed edges that enclosed small areas. The complexity of this level is not known, and our ability to predict performance by using simple rules implies that the visual system does not use complicated ones (in the case of untrained observers and stimulus uncertainty). The results of our psychophysical experiments (described in Section 5) show that asymmetry also occurs in detection tasks without stimulus uncertainty (but with spatial uncertainty) and when one is using trained observers.

We suspect that the decision strategy used will be particularly important in discrimination tasks of textures having similar power spectra, such as pair 13 of Fig. 5. Although the two elements of this pair have the same energies at all spatial frequencies, the distribution across orientation is different. At low spatial frequencies, one of them (arrow) has a relatively small orientation variance, whereas the other (triangle) has a pronounced variance, thus producing noisy filtered maps when the figures are randomly oriented. This situation can be enhanced by using the classical triangle-arrow pair with equal sides,³⁸ producing an even smaller orientation (and thus spatial) variance for the arrows. The result is a much better discrimination rate (41% for arrows and 58% for triangles as foreground), in agreement with psychophysical observations.^{12,39} Nothdurft⁴⁰ observed a reduction in discrimination of these textures when spatial jitter is increased, thus providing more evidence for the role of spatial variability in this task.

The overall discrimination performance of our model is determined by the noise variance (V_n), which may depend on internal noise and mask noise. However, in the cases examined here, the main limiting factor is the texture spatial variance. An absolute limit on model performance was obtained by testing the model with $V_n = 0$, resulting in some improvement for most pairs (0 to 24%, with an average of 6.4% and a standard deviation and 6.7%). This improved performance is equal, on the average, to human performance obtained at long stimulus durations of 167 msec in the experiments reported by Gurnsey and Browse.¹

Finally, asymmetry is not necessarily a result of having unequal background and foreground areas. We predict that asymmetry depends on foreground positional uncertainty (number of possible locations) and not so much on area ratio. However, within the framework of the model, the foreground area can be important in two ways: relative to the smoothing operator size and relative to the size of clusters formed by background variability. These two factors should be examined more carefully, both theoretically and experimentally.

5. PSYCHOPHYSICAL STUDY OF ORIENTATION VARIABILITY

In order to test directly the dependence of texture segmentation on orientation variability, we conducted psychophysical experiments. These experiments attempt to find a correspondence between texture-orientation variability and performance asymmetry. It should be noted that these experiments were not modeled by the previously described model, because of the different experimental paradigm used. Although they too could be modeled, we found the extensive

data provided by Gurnsey and Browse to be sufficient in supporting our model. These experiments rather were performed only to confirm the importance of orientation variability.

A. Experiment Design

The experiments were divided into three subexperiments, each of which had different orientation requirements. The first subexperiment employed elements that had identical orientation (0°), the second had elements having two possible orientations (0° and 180°), and the third used randomly oriented elements (12 orientations at 30° intervals). Figure 6 shows examples of typical stimuli. The motivation for the design of these subexperiments was to see how asymmetry changes from one subexperiment to another. Specifically, if asymmetry increased with increasing element-orientation randomization, then orientation variability components are critical characteristics in the discrimination task. However, we suspected that if we use energy measurements based on Gabor-like filters, then using elements having two orientations, 180° apart, would not change the performance levels compared with those for aligned elements. That is to say, because Γ 's have redundant Gabor energies every 180° , performance levels in the first subexperiment should be identical to those in the second.

B. Methods

Five observers, two of whom were the authors, participated in the experiments. The others were high school students who were paid to participate and were not aware of the purpose of the experiments. A Hewlett-Packard Model 1310B oscilloscope (P31 phosphor), driven by custom-designed hardware⁴¹ allowing for real-time control of the stimulus properties, was used in an isolated (dark) environment. Screen resolution was 1024×1024 pixels. The experiments and the graphic device were controlled by a Sun Model 3/160 workstation. The observer was presented a stimulus that

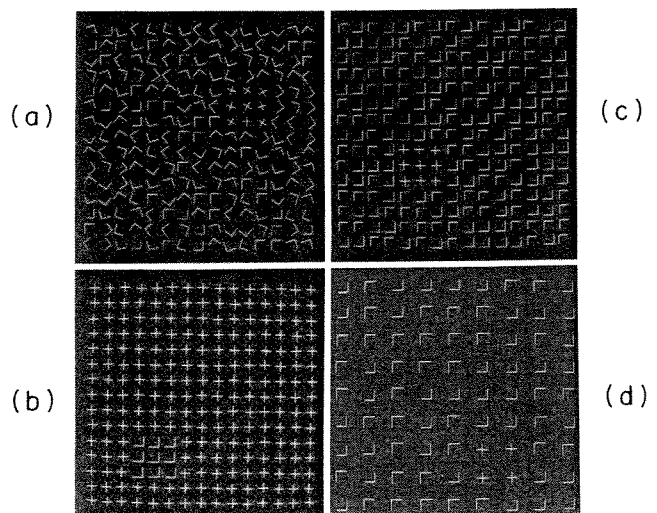


Fig. 6. Typical stimuli of psychophysical experiments: (a) has elements with random orientations, (b) has aligned elements, and (c) has elements with two orientations; (a)–(c) have dense spacing and (d) has sparsely spaced elements of two orientations.

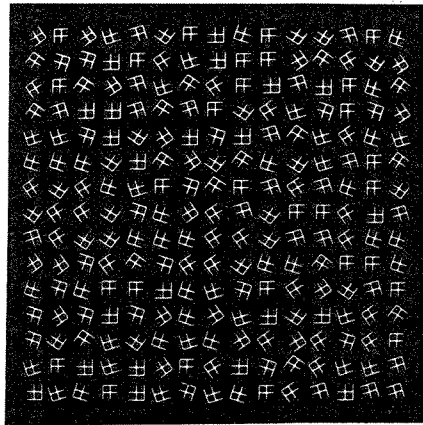
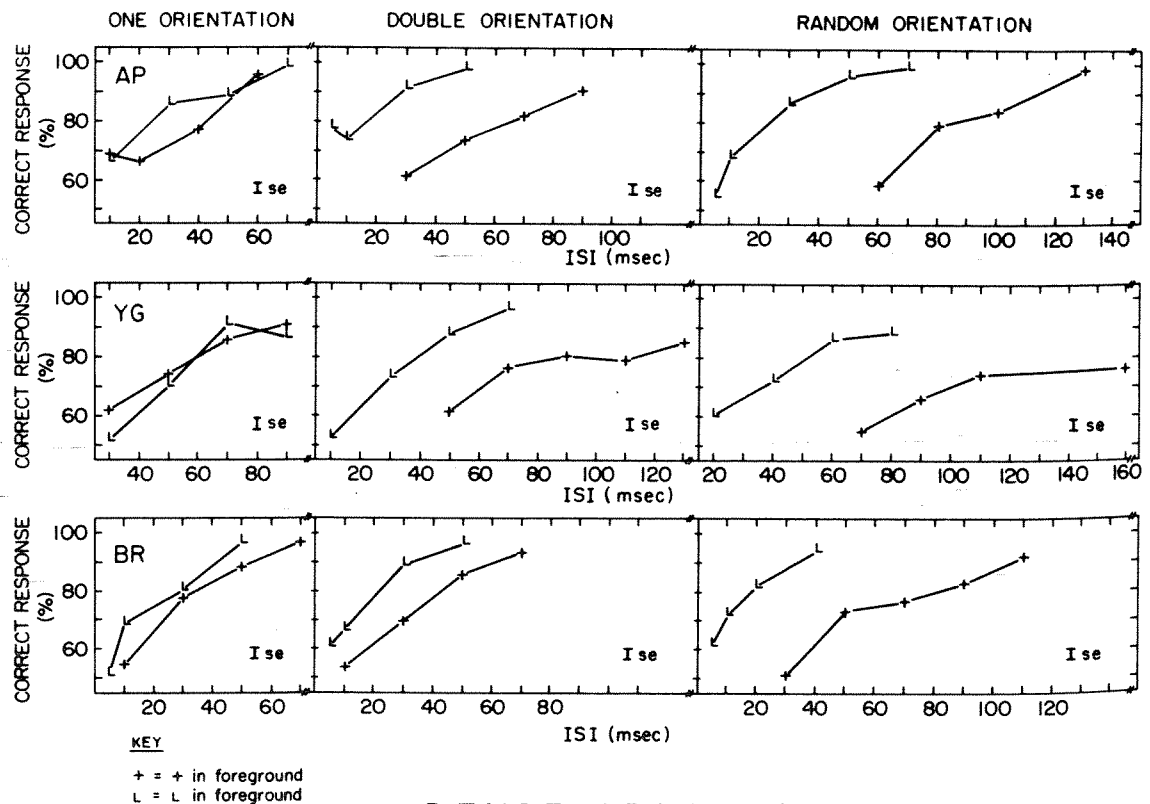


Fig. 7. Typical mask having elements comprising a $+/ \Gamma$ combination shown with randomly generated orientations.

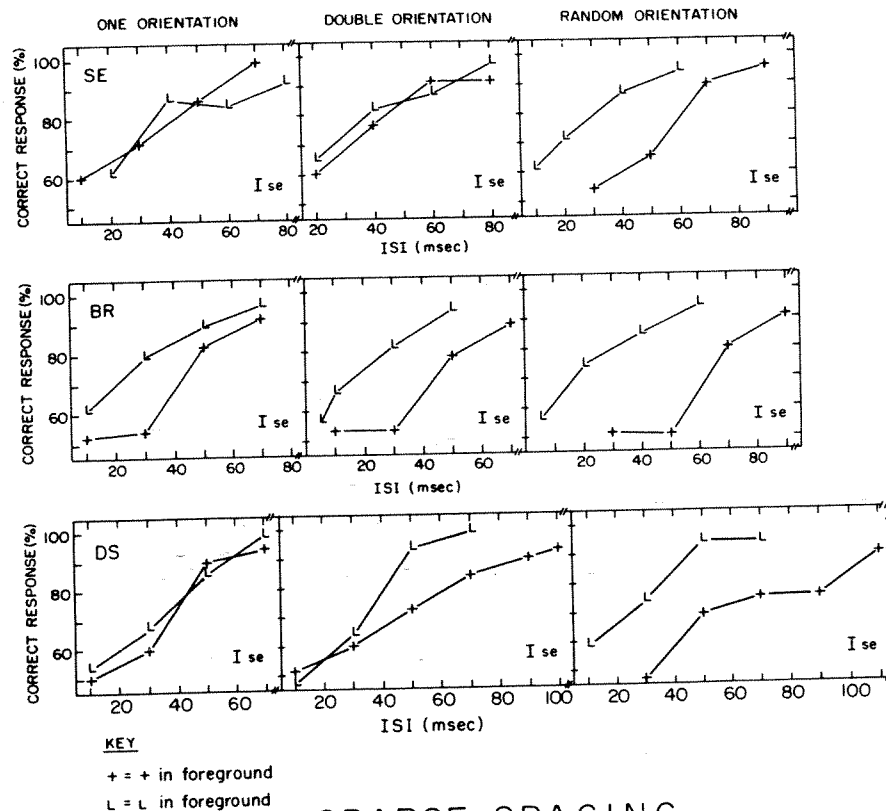
contains a background and an intermittently shown foreground region followed by a mask (see Fig. 7). Consequently the observer had a simple detection task. The stimulus was composed of a ground, a 15×15 matrix of elements, and a disparate region, a 3×3 matrix randomly situated over the ground. The foreground placement had been defined not to overlie the center fixation point so the observer could not simply identify the target by merely studying the center area. Also, the foreground did not overlie the bordering two

elements of the ground so as to stay within the foveal area. Each element was placed within a 64×64 pixel field (100 in sparse spacing conditions). With sparse spacing, the ground was a 9×9 matrix of elements, and a 2×2 matrix was used for the disparate region. $+$'s and Γ 's were used with line segments of each element being 40 pixels in length. In turn, each element field was placed adjacent to another in the stimulus field. Also, each element was positioned randomly (jitter) about the center of its space, with as many as pixels in any direction. This unalignment was implemented to reduce the potential assistance of global alignment in the detection task. The observer was positioned 170 cm away from the screen, subtending a 20.2° arc to each element segment. The side of the ground matrix subtended an 8.1° arc. During the actual experiment, the fixation mark was shown to center the observer's attention followed by the stimulus (shown for a duration of 10 msec = 1 frame) with or without the disparate region (target). Then after an interstimulus interval (isi) of 5 to 160 msec, a mask composed of randomly oriented elements ($+/ \Gamma$ combination with 12 orientations at 30° intervals) was shown (100 msec). Note that the mask's matrix size was identical to that of the stimulus, positionally aligned to overlap except for the 6 pixel jitter. The disparate region was presented only 50% of the time, and each observer was asked to detect it by using the keyboard to answer yes (1) or no (0). Experiments were conducted in blocks of 50 trials, with each block having exactly the same conditions (i.e., isi and elements remain



DENSE SPACING

Fig. 8. Psychophysical results of three observers, AP, YG, and BR, for densely spaced stimuli having one orientation, two orientations, and random orientations. Asymmetry is marked by the spaces between the curves. Notice that asymmetry increases for textures having more orientations. The average standard error is depicted at the lower right-hand portion of the graph box.



SPARSE SPACING

Fig. 9. Psychophysical results of three observers, SE, BR, and DS, for sparsely spaced stimuli having one orientation, two orientations, and random orientations. Notice that there exists great asymmetry for random orientation; however, there is little increase in asymmetry from one to two orientations.

in their respective background-foreground regions, but the foreground was in different positions). Subjects were tested with both '+'s and 'L's in the disparate region. Before each experiment, observers were allowed to see the stimulus in order to understand the task and were aware of the ISI. Detection rates were calculated as the average of the performances to detect both conditions (with or without target) so as to eliminate any subjective preference to one of the conditions. Each observer was limited to 1 h of experimentation to offset the effects of fatigue. Also, observers were allowed to practice for extended periods (up to weeks) so as to acclimate themselves to the task at hand and, more importantly, learn the task to the point that they can reproduce similar results.

C. Results

The psychophysical data clearly show that, as the elements have more variability in orientation, asymmetry increases. Figure 8 shows the data with asymmetry marked as the horizontal spaces between the curves. In the first subexperiment (one orientation), discrimination for both elements in the foreground produces similar performance levels, showing that the task is symmetric. In the second and third parts (double and random orientation), asymmetry becomes greater and greater, indicating the importance of orientation variability. Also, note that, contrary to our prediction concerning the second subexperiment of two orientations, asymmetry does increase.

However, in studying the stimulus, we decided that perhaps there were some additional energy variabilities caused by the dense spacing of elements that was affecting the detection process. Specifically, for the two-orientation subexperiment, line segments of adjacent 'L's tended to lie closer to one another at times in comparison with the first subexperiment (one alignment) [see Figs. 6(b) and 6(c)]. This would cause greater variability in energies, leading to greater asymmetry than that of aligned 'L's. In order to remove this unwanted feature and to isolate orientation variables, we increased the spacing between the elements, placing the elements in a larger 100×100 pixel space [see Fig. 6(d)]. This would remove much of the accidental merging of elements and reduce undesired variability. Hence we conducted an experiment identical to the one mentioned above with the exception of larger spacing, defining a 9×9 matrix of elements for the ground and a 2×2 matrix target. Figure 6(d) is one example of such a stimulus.

The results are similar to the previous experiments with the exception that, as expected, asymmetry does not increase measurably from the first subexperiment to the second (see Fig. 9). This supports the idea that orientation cells detect little additional variability, causing performance levels to remain the same. The small increase in asymmetry can be accounted for by the effects of positional variability mentioned in Section 1, but they are shown here to be negligible (in the case of large spacing) compared with the effects of orientation variability. Also, note that one observer ex-

hibited asymmetry for the first subexperiment (one orientation) of sparse spacing (see results for observer BR in Fig. 9). However, this does not affect the validity of the conclusions.

6. CONCLUSION

In conclusion, we have shown experimentally and theoretically that orientation variability can account for performance asymmetries in texture-discrimination tasks. Our experiments show conclusively that, for textures having randomly oriented elements, those elements with relatively greater orientation variability are more distracting in the background, resulting in decreased discrimination. Hence asymmetry is the result of background-foreground elements with different variability profiles. Furthermore, it should be noted that the essence of this model is not to emphasize the importance of variability across the orientation spectrum but rather to show how noise characteristics of a stimulus are the critical contributors in testing whether discrimination is possible and to what degree. Orientation variability is hence merely one kind of noise, and, although it has been found to be important in the examples mentioned above, other noise characteristics such as interelement characteristics and those concerning the vertical and tilted lines described by Treisman^{6,7} can also be important noise characteristics of a stimulus and therefore must be considered.

In spite of the different noise characteristics, we found that orientation variability is a major determinant of discrimination performance. In fact, using a model based on orientation variability, we found a remarkable correlation with the psychophysical data of Gurnsey and Browse.¹ The model's success shows that sensory processing underlying texture discrimination can be understood strictly in terms of local computations, that is, no global calculations need be invoked. Previously models found local computations to be insufficient and therefore included global calculations, thus making the problem more complex and difficult.⁹ In our model, an inevitable global component enters only at the decision-making stage, i.e., when a global maximum of local differences must be found. This stage should be task dependent, since the decision rule that we used for the spatial our-alternative forced-choice task is not applicable in detection tasks. Detection experiments (Section 5) contain no-target trials that produce global maxima in the same way as target trials. The solution suggested by standard detection theories³⁴ is to use a threshold in order to separate true maxima from false ones. The success of this thresholding operation in distinguishing between trials with and without foreground depends on the magnitude of the F-B difference signal relative to the background noise level.

On a biological level, the implications of such a model are twofold. First, because of the highly satisfactory performance of a model based on orientation filters and local computations (resembling orientation-selective cells in the visual cortex), texture segmentation would seem to be processed at a low level in the visual system. Second, this low-level processing implicates preattentive vision, a stage that deals with all parts of the visual field in parallel. Third, because we find orientation variability to be important in texture discrimination, both theoretically and psychophysically, perhaps there exists a cortical stage of processing that

connects nearby orientation cells and consequently can detect local activity differences.

ACKNOWLEDGMENTS

We thank Jochen Braun for his comments, Anne Reich for her general assistance, and Yehuda Barbut for figure preparation. This study was supported by the Basic Research Foundation administered by the Israel Academy of Sciences and Humanities.

All correspondence should be addressed to Dov Sagi.

REFERENCES

1. R. Gurnsey and R. Browse, "Micropattern properties and presentation conditions influencing visual texture discrimination," *Percept. Psychophys.* **41**, 239-252 (1987).
2. J. Beck, "Textural segmentation," in *Organization and Representation in Perception*, J. Beck, ed. (Erlbaum, Hillsdale, N.J., 1982).
3. H. C. Nothdurft, "Sensitivity for structure gradient in texture discrimination task," *Vision Res.* **25**, 1957-1968 (1985).
4. D. Sagi and B. Julesz, "'Where' and 'what' in vision," *Science* **228**, 1217-1219 (1985).
5. D. Sagi and B. Julesz, "Short-range limitation on detection of feature differences," *Spatial Vision* **2**, 39-49 (1987).
6. A. Treisman, "Preattentive processing in vision," *Comput. Vision Graphics Image Process.* **31**, 156-177 (1985).
7. A. Treisman, "Features and objects in visual processing," *Sci. Am.* **255**, 106-125 (1986).
8. R. Gurnsey and R. Browse, "Aspects of visual texture discrimination," in *Computational Processes in Human Vision: An Interdisciplinary Perspective*, Z. Pylyshyn, ed. (Ablex, Norwood, N.J., 1988).
9. A. Treisman and S. Gormican, "Feature analysis in early vision: evidence from search asymmetries," *Psychol. Rev.* **95**, 15-48 (1988).
10. A. Treisman and J. Souther, "Search asymmetry: a diagnostic for preattentive processing of separable features," *J. Exp. Psychol.* **114**, 285-310 (1985).
11. B. Julesz, "A brief outline of the texton theory of human vision," *Trends Neurosci.* **7**, 41-45 (1984).
12. B. Julesz, "Texton gradients: the texton theory revisited," *Biol. Cybern.* **54**, 245-251 (1986).
13. I. Fogel and D. Sagi, "Gabor filters as texture discriminator," *Biol. Cybern.* **61**, 103-113 (1989).
14. J. Beck, A. Sutter, and R. Ivry, "Spatial frequency channels and perceptual grouping in texture segregation," *Comput. Vision Graphics Image Process.* **37**, 299-325 (1987).
15. T. M. Caelli, "Three processing characteristics of visual texture segmentation," *Spatial Vision* **1**, 19-30 (1985).
16. J. D. Daugman and D. M. Kammen, "Pure orientation filtering: a scale-invariant image-processing tool for perception research and data compression," *Behav. Res. Meth. Instrum. Comput.* **18**, 559-564 (1986).
17. M. R. Turner, "Texture discrimination by Gabor functions," *Biol. Cybern.* **55**, 71-82 (1986).
18. B. Julesz, "Visual pattern discrimination," *IRE Trans. Inf. Theory* **8**, 84-92 (1962).
19. B. Julesz, H. L. Frisch, E. N. Gilbert, and L. A. Shepp, "Inability of humans to discriminate between visual textures that agree in second-order statistics," *Biol. Cybern.* **31**, 137-140 (1973).
20. D. A. Pollen and S. F. Ronner, "Visual cortical neurons as localized spatial frequency filters," *IEEE Trans. Syst. Man Cybern.* **SMC-13**, 907-916 (1983).
21. F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," *J. Physiol. (London)* **197**, 551-566 (1968).
22. J. Daugman, "Two dimensional spectral analysis of cortical receptive field profiles," *Vision Res.* **25**, 671-684 (1980).