Pergamon

PII: S0042-6989(96)00235-0

Vision Res., Vol. 37, No. 21, pp. 2993–3007, 1997

Printed in Great Britain
0042-6989/97 $17.00 + 0.00

The Orientational Resolution of Human Texture Perception

D. R. T. KEEBLE,*†‡§ F. A. A. KINGDOM,* M. J. MORGAN†

Received 15 December 1995; in revised form 1 August 1996

A major determinant of human texture segregation and discrimination is the orientational content of the stimuli used. We have investigated the ability of observers to resolve features defined in the orientation domain in a variety of textures. It was found that features had to be separated by at least 13 deg for subjects to discriminate orientationally bimodal textures from same-variance unimodal textures. For larger separations, the determinant of performance was the magnitude of the central "dip" in the probability density functions determining the bimodal textures. Resolution performance can be modelled by assuming that a filtering process over orientation demodulates the central dip in the bimodal texture and that discrimination depends on criterion depth in the resulting function. Such modelling produces relatively broad estimates of the bandwidth ranging from about 10–20 deg. Performance was similar for both line and Gabor micropattern stimuli.

Texture Psychophysics Orientation Filtering Resolution

INTRODUCTION

The textural content of an image region is a cue that the visual system can use for segmentation and identification (Beck, 1966; Julesz, 1962), as well as surface orientation estimation (Gibson, 1950). It is widely believed that the orientational properties of a texture are particularly important in this regard (Beck *et al.*, 1983). The majority of recent models (e.g. Landy & Bergen, 1991; Malik & Perona, 1990; Graham *et al.*, 1992) of human texture perception use arrays of oriented filters as their input stage, in accord with the known ubiquity of orientationally selective neurons in V1 (Hubel & Wiesel, 1959). Thus, it is important to understand how good the visual system is at detecting differences in the orientational content of textures, for such knowledge places constraints upon possible models of texture processing.

We have previously shown that human performance in the discrimination of textures with threshold non-random orientational content from orientationally random (isotropic) textures can be modelled well by assuming that the visual system has access to the output of a global filtering operation over orientation (Keeble *et al.*, 1995a, b). The filter that emerged from the modelling process was broad, having an average half-height full-width of 34 deg. These experiments were essentially concerned with the detection of any kind of non-random orientational content, where the content could be multimodal in the orientational domain. This loosely corresponds to the concept of contrast sensitivity in the luminance domain (Campbell & Robson, 1968). A related but distinct issue is the question of how good observers are at discriminating between textures with different kinds of orientationally supra-threshold content. That is, if a texture clearly contains one or more dominant directions, how well can we assess the properties of the orientational content?

A fundamental capability of the visual system across several domains is that of resolution, or the ability to decide whether there is one feature or many in some image region, and the level of performance in distinguishing two spots of light from one spot of light is perhaps the paradigmatic example of a resolution task. The standard measures of acuity used by optometrists, such as the Snellen eye chart, Landolt C-ring etc. are also examples of resolution tasks, as they are based on performance in distinguishing proximate spatial features. A measure termed the resolution limit or limiting frequency has been used in vision science to quantify the resolving power of the human observer (Bennett & Rabbetts, 1984). This measure is essentially the spatial

*McGill Vision Research, Department of Ophthalmology, McGill University, 687 Pine Avenue West, H4-14, Montréal, Québec, Canada H3A 1A1.

†Institute of Ophthalmology, University of London, Bath Street, London EC1V 9EL, U.K.

‡To whom all correspondence should be addressed [*Email*: dkeeble@jiffy.vision.mcgill.ca].

§Present address: Information Science Research Laboratory, NTT Basic Research Labs, 3-1 Morinosato-Wakamiya, Atsugi-shi, Kanagawa 243-01 Japan.

frequency at which observers can only just detect a sine-wave grating at maximum contrast. Although it is possible to estimate an analogous textural resolution limit from our previous work (see Discussion), this estimate only relates to performance at threshold modulation.

In order to assess the resolution capabilities of the texture mechanism of the human visual system we use an analogue of the two-line resolution task in the luminance domain (e.g. Levi & Klein, 1990). The task of the observer in that experiment is to distinguish two spatially proximate lines from just one line. In our experiment, observers must discriminate between textures comprising oriented micropatterns which are globally either orientationally unimodal or bimodal. In other words, the task is to assess whether there are one or two dominant orientations in the texture. We control the orientational content by constructing the textures with micropatterns (lines or Gabor patches), with orientations drawn from probability density functions (pdfs), defined over orientation, as in our previous work (Keeble *et al.*, 1995a, b). The pdfs could either be Gaussian in form (unimodal), or be two Gaussians added together but with peaks at different angles (bimodal). In the limit, as the standard deviations of these sum-of-offset-Gaussians (SOOG) pdfs approach zero the texture will comprise lines with just two orientations. An example of a SOOG pdf is shown in Fig. 1. Using this flexible way of defining the bimodal texture allows us to assess resolution performance with different ranges of orientation (i.e. orientation variance) in the stimulus, and different angular separations between the orientation maxima. We ensure that we are employing a genuine resolution task, and not just a variance discrimination task, by matching the orientation variance between the bimodal texture and the unimodal comparison texture.

These experiments address a number of specific issues. Is it possible to model resolution performance using a linear filtering operation over orientation, as proved possible for the detection of textural anisotropy (Keeble *et al.*, 1995a, b)? If so, will the filter be of a similar width? What will be the relevant parameters of the bimodal pdf for performance? What is the absolute limit of resolution performance in the optimal stimulus? How does the nature of the oriented micropatterns affect performance? The main result of this paper is to show that simple linear modelling can account for the data well, assuming that the relevant feature is the amplitude of the "dip" in the middle of the bimodal distribution.

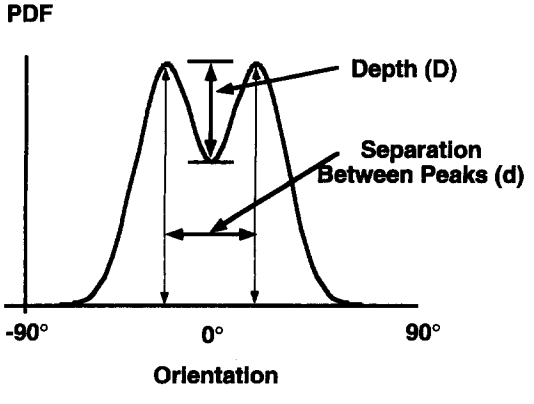

FIGURE 1. The pdf of a SOOG. The depth of the central dip (D) and the angular separation between the peaks (d) are shown.

METHODS

The methods used to conduct this experiment were in general similar to those used in Keeble *et al.* (1995a, b), so we present an abbreviated description here.

Display

Stimuli were produced with the use of Apple Macintosh computers (IIcx and Quadra 650) and were displayed on two different Macintosh 13" colour displays, each of which was appropriately gamma-corrected. The stimuli consisted of arrays of anti-aliased bright white lines on a dark background, or arrays of Gabor patches on a grey background. Gabor patches are sinusoidal gratings isotropically windowed by a Gaussian function. The patches we used were in cosine phase, and the ratio of the amplitude of the underlying sinusoid to the mean luminance of the screen about which it modulates was 0.98. The mean luminance was 48 cd m^{-2}. The standard deviation of the sinusoid was 8.8 min arc of visual angle, and three different peak frequencies, 2.0, 4.1 and 8.2 c deg^{-1}, were used. The length of the lines was 33 min arc and their width was 2.2 min arc. Pixel resolution at the 57 cm viewing distance was 27.2 pixels deg^{-1}. The experiments were performed under room lighting. Two durations of stimulus exposure were used for most of the experiments: 105 msec and 1000 msec. Detailed inspection of the stimuli was prevented by the shorter duration, and the longer duration allowed us to investigate how resolution performance improved with time. At the longer duration subjects were allowed to make eye movements, if they wished.

Stimuli and procedure

A two-alternative-forced-choice (2AFC) task was used. Two spatially adjacent texture arrays were shown to the subjects. One array was orientationally bimodal, whilst the other was orientationally unimodal. Subjects were required to select the array that was bimodal. Feedback for incorrect responses was provided in the form of a beep. Equal numbers of presentations had the bimodal array on each side. Examples of the stimuli used, together with the pdfs used to generate them are shown in Figs 2 and 3.

Each array was 10.6×10.6 deg in size, the two arrays being horizontally separated by 44 min arc. In Experiments 1 and 2 there were 500 randomly positioned lines in each of the two arrays, whereas in Experiment 3 there were 196 micropatterns positioned on a 14×14 grid with uniform random positional jitter between ± 13.2 min arc added in each dimension to hinder the

occurrence of collinear structures. Grid stimuli were utilised so as to prevent the Gabor patches overlapping, which would have reduced contrast and thus the visibility and salience of the stimuli that would have been possible to use. This is because the contrasts of overlapping Gabor patches are combined additively. When line stimuli overlapped the pixel luminances at the intersection were not added.

In Experiment 1, on each presentation the bimodal pdf was a sum of offset Gaussians (SOOG). By "offset", we mean that the maxima of the Gaussians were separated by an angle that we shall refer to as "*d*". The comparison unimodal pdf was Gaussian. We used the angular separation of the peaks in the SOOG, *d*, as the independent variable for each psychometric function, by analogy with the luminance domain, where resolution, or acuity, is generally measured as a function of either the separation between spatial features, or the spatial frequency (Bennett & Rabbetts, 1984). It became obvious in trial observations that if the standard deviation of the unimodal distribution, σ_U, and the individual Gaussians, σ_G, in the SOOG were kept constant then the concomitant difference between the overall orientational variances of the two arrays became an extremely powerful cue. In a similar task, Dakin (1994) has shown that subject performance is predicted extremely well by a variance discrimination model. In order to prevent the experiment from being an orientation variance discrimination task, and thus to ensure that it was a genuine test of resolution ability, we required that on each trial the orientational standard deviation be the same for the unimodal and bimodal pdfs. Thus, as d increased the standard deviation of the underlying Gaussians of the SOOG was decreased, in order to keep the overall standard deviation of the pdf constant. In Experiment 1, where SOOGs were used, the orientational variance over each psychometric function was thus constant. Figure 2 shows two examples of these stimuli and their corresponding pdfs at different standard deviations.

In Experiments 2 and 3 the bimodal distribution was just two discrete orientations—that is, the pdf was a normalized sum of two impulse functions. In this case the comparison pdf was again Gaussian, but here increasing *d* in the bimodal pdf increases the standard deviation, so the standard deviation of the Gaussian must be increased to match. Examples of these stimuli for different spatial arrangements of the lines and for different micropatterns are displayed in Fig. 3.

Although the example stimuli shown each have mean orientations which are vertical, in the actual stimuli the orientational phase of the pdfs was random from trial to trial. This arrangement ensures that subjects cannot perform the task by attending to the orientation strength in one prespecified direction. In a given stimulus the mean orientations of the unimodal and bimodal arrays were the same. It should be noted that in Experiment 1, in particular, the use of continuous pdfs (i.e. the existence of a range of orientations), means that the task must be achieved by some spatially integrative mechanism, taking information from many first-order units, rendering this a true texture task, and not simply by a local detection of orientation properties (i.e. just using the information from a couple of micropatterns).

In each experiment the percent correct was plotted as a function of the angular distance, *d*, between the maxima of the pdf for the bimodal texture. In Experiment 1, six equispaced values of *d* were used, with appropriate values being determined by pilot experiments. It should be pointed out that in this case values of *d* have a maximum permissible value of twice the overall standard deviation of the orientations in either of the arrays. This is because in the limit as *d* increases the individual Gaussians in the SOOG go to impulse functions as the bimodal pdf changes shape in order to keep the orientational variance constant. Beyond this point *d* cannot increase without producing a corresponding increase in the standard deviation, thus violating the previously discussed requirement of equal variance between the bimodal and unimodal pdfs in a given stimulus. In passing, we note that the overall standard deviation of a SOOG, σ_{SOOG}, is given by:

$$\sigma_{SOOG} = \sqrt{(\sigma_G^2 + d^2/4)}, \tag{1}$$

where σ_G is the standard deviation of the component Gaussians, and *d* is the offset between the peaks. All the different standard deviation conditions were run together in the same block for this experiment, with different exposure durations being in different blocks. In Experiments 2 and 3, eight equispaced values of *d* from 5 to 40 deg were used for each psychometric function. Different micropattern, spatial arrangement of micropatterns (i.e. random or grid) and duration conditions were run in different blocks. In all experiments 40 observations were collected for each point on the psychometric function. The sole exception to this was observer FK in Experiment 3, where we took 120 observations per data point in order to verify that the small decrease in threshold between 4.1 and 8.2 c deg^{-1} in the 1000 msec condition was real. For Experiment 1, observers DK and FK collected the data in 10 separate blocks for each duration condition, whereas PL spread the collection over 20 blocks. The data for each condition for the other experiments was taken in two blocks, except for the above mentioned exception, which was taken in six blocks. The subjects all performed at least one practice block for every condition and experiment in order to familiarize themselves with the task.

Weibull functions (Weibull, 1951) were fitted to the psychometric functions. These are given by:

$$P(d) = 1 - \frac{1}{2}\exp[-(d/\alpha)^{\beta}], \tag{2}$$

where α is a threshold parameter (corresponding to the 82% correct point), β is the slope parameter and *d* is the separation between the peaks. For all our data we take α as the threshold. The fitting was done by a maximum likelihood method which produced 67% confidence

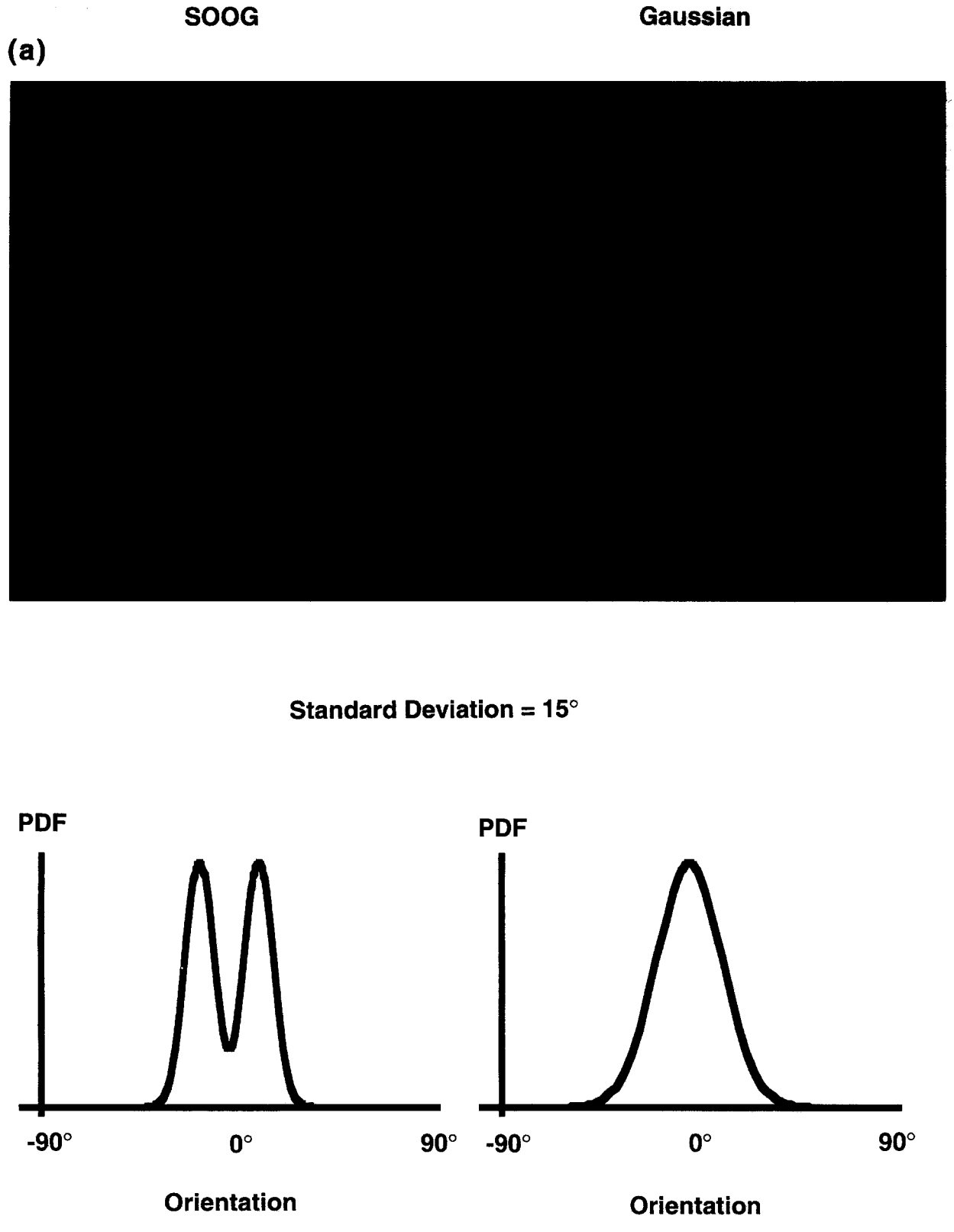

FIGURE 2(a). *Caption on facing page.*

intervals and a χ^2 criterion for acceptance or rejection of the fit at the $P < 0.05$ level. Thresholds were discarded in Experiment 1 if they corresponded to mathematically impossible stimuli, i.e. if threshold d was larger than was possible to produce for a SOOG at this overall variance. Otherwise, only two psychometric functions in Experiment 1 failed the goodness-of-fit test. These were for PL at 105 msec, and DK at 1000 msec, both at standard deviations of 25 deg. Two failures of a 95% test out of 37 data points is about what one would expect by chance,

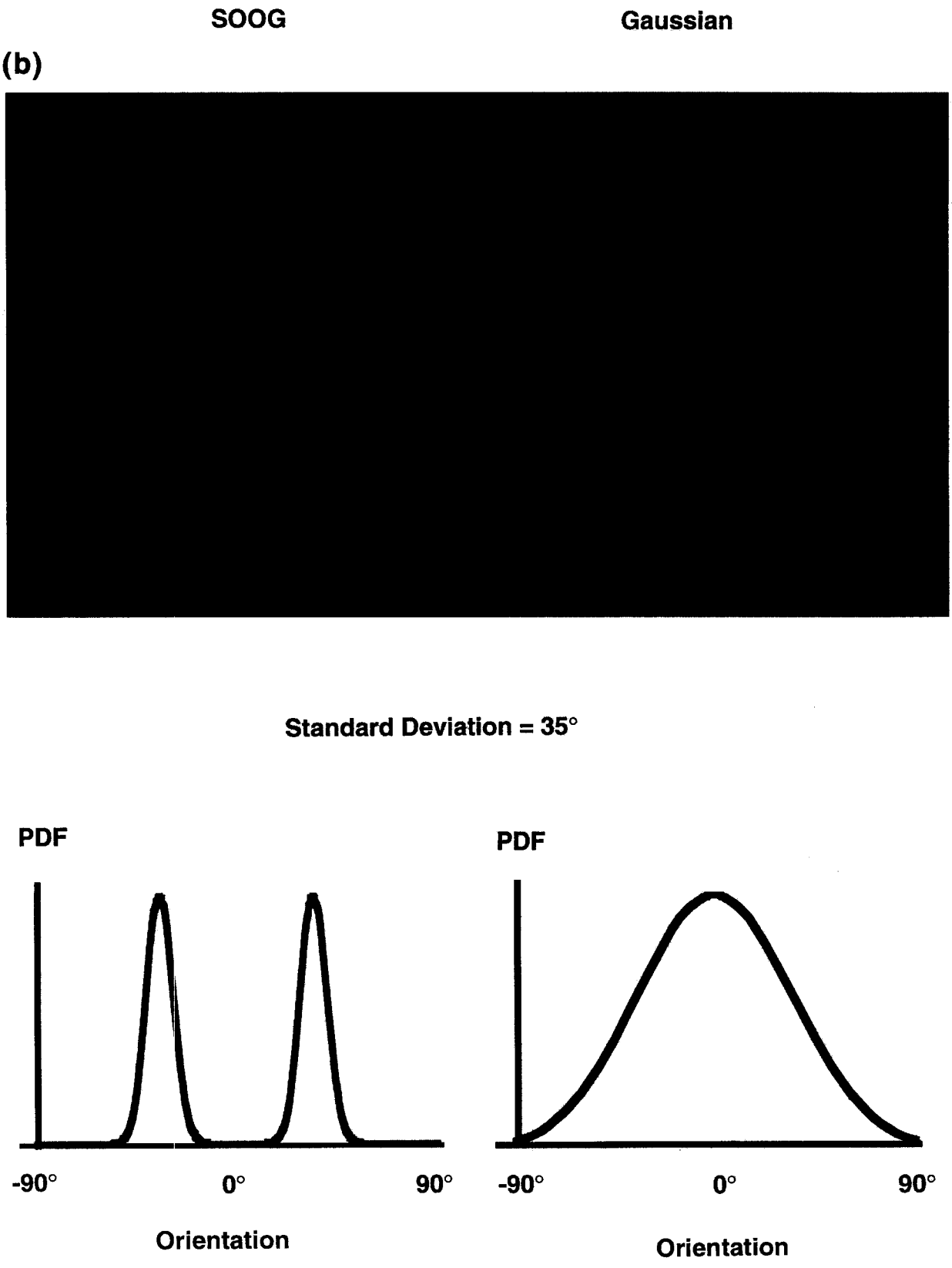

FIGURE 2. Two examples of the SOOG stimuli used in Experiment 1, together with their generating pdfs. In (a) the overall standard deviation is 15 deg, whilst in (b) it is 35 deg.

and the thresholds produced by these data points are very much in accord with the pattern of change of threshold with orientation variance, so we have chosen to plot these two values and include them in the analysis and modelling in the next section.

Subjects

Three subjects performed these experiments. Two (DK and FK) were authors, and the third (PL) was a graduate student. All subjects were experienced psychophysical observers.

FIGURE 3(a,b). *Caption on facing page.*

RESULTS AND ANALYSIS

Experiment 1—Resolution as a function of orientation variance

In Experiment 1 the ability of observers to resolve textural features at varying overall orientational variances is measured. If there is some fundamental limit to orientational resolution, it should manifest itself to a different extent over different ranges of orientation. Each threshold is the angular separation, d, between the maxima of the SOOG at which criterion performance is achieved. The results for three observers are plotted in Fig. 4. Not surprisingly, as the overall variance increases the separation between the peaks required to do the task

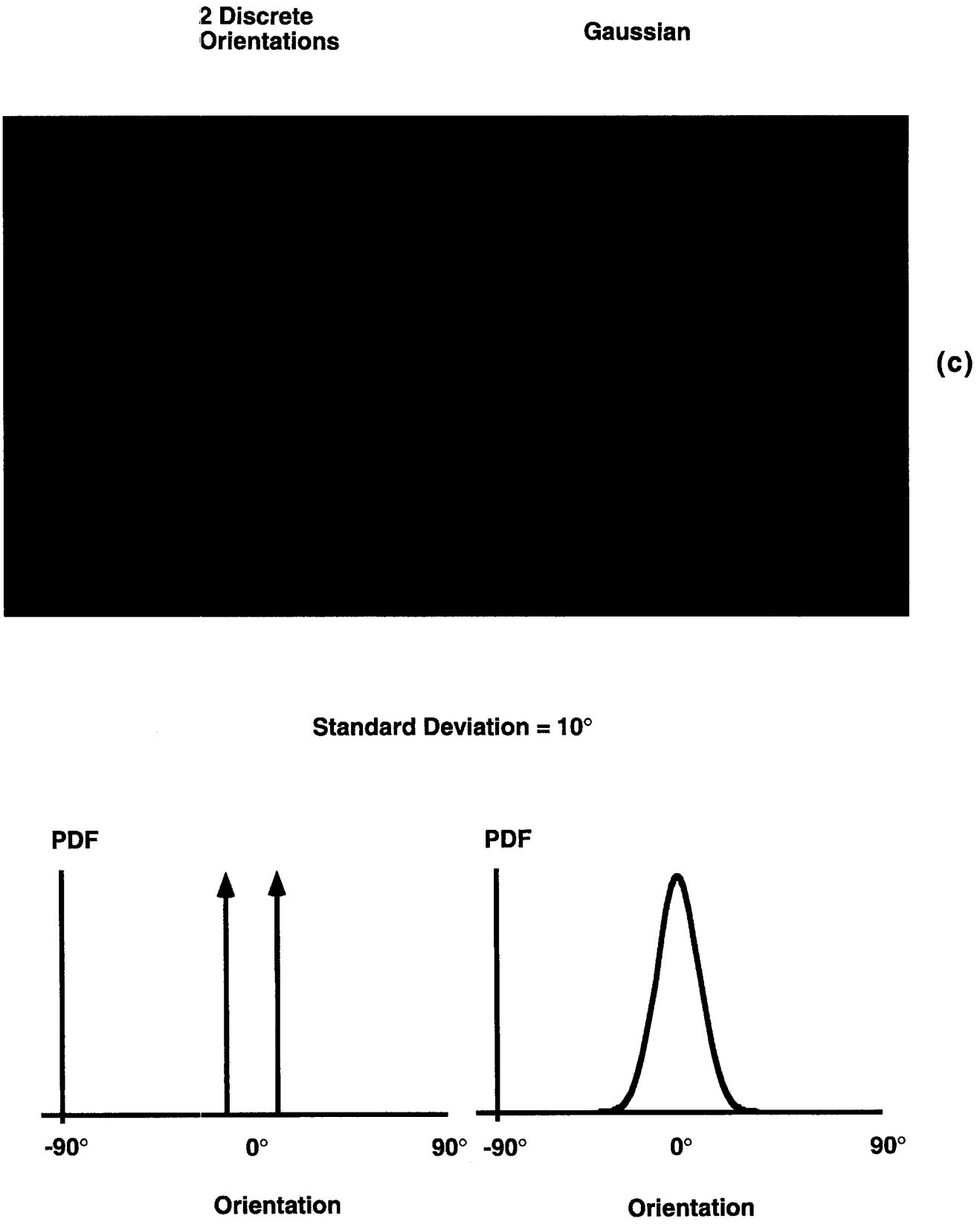

FIGURE 3. Examples of the two discrete orientation stimuli used in Experiments 2 and 3. In each case the overall orientational standard deviation is 10 deg. (a) The lines are placed randomly. (b) Gabor micropatterns of peak spatial frequency 4.1 c deg^{-1} are placed on a jittered grid. (c) Line elements are placed on a jittered grid.

increases. This increase is approximately linear. It should be born in mind, however, that the thresholds are effectively constrained to fall within the straight lines shown in the diagram. The upper line ($d = 2\sigma_{SOOG} = 2\sigma_U$) corresponds to the peak separation at which the bimodal distribution is two impulse functions, i.e. the maximal

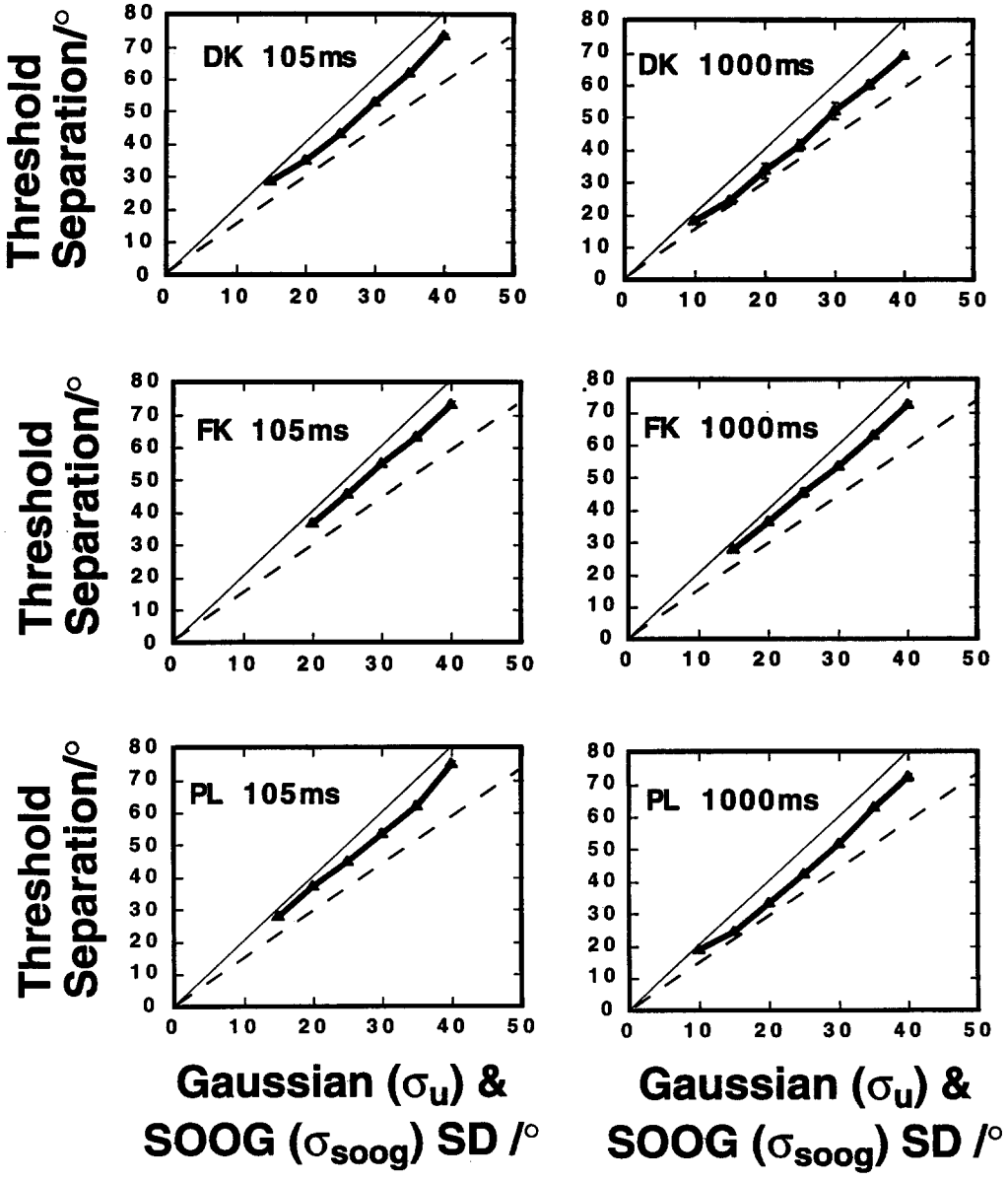

FIGURE 4. Threshold peak separation for resolution performance as a function of overall orientation standard deviation for Experiment 1. See text for explanation of the upper and lower limiting lines.

separation possible when the bimodal and unimodal variances are constrained to be the same. This follows from equation (1). On the other hand, the lower line ($d = \sqrt{2\sigma_{SOOG}} = \sqrt{2\sigma_U}$) is the separation between the individual Gaussians where the central dip in the pdf disappears. We note that at the lowest standard deviations it was not possible to obtain thresholds, as evidenced by the way the data meets the $d = 2\sigma_{SOOG}$ line at low σ_{SOOG} in each case. It is conceivable, but not likely, that discrimination could occur between two unimodal pdfs with the same variance and similar shapes, but evidently this is not possible in this case. So it is clear that although plotting the data in this fashion shows that the task is possible under these conditions, it is compressing the possible data space into a wedge-shaped region. This compression makes it difficult to evaluate different models for the subjects' performance. Levi and Klein (1990) found a similar linear relation at high standard deviations in their two-bar luminance resolution experiment. Because they had no equivalent to our equal variance constraint, however, they were able to evaluate thresholds at small standard deviations, allowing them to find a flat region of the curve. Whatever the merits of this approach in the luminance domain, it was clear from our pilot observations that relaxation of the equal variance constraint would have rendered the bimodality of the stimulus irrelevant to the task.

Because this conventional manner of plotting the results proved not to be particularly enlightening, we sought another metric more directly related to the bimodality of the stimulus, which is illustrated in Fig. 1. Instead of taking the separation between the peaks, we evaluated the depth, D, of the pdf. For each orientational variance we took the pdfs at threshold performance, and

evaluated the depth of the central dip. In this fashion we replot the results in Fig. 5. The errors bars shown are the mean transformed endpoints of the 67% confidence intervals of the threshold separations. There are two salient features of this representation of observer performance. First, as the orientational variance increases, depth, D, at threshold d becomes approximately constant. Second, at smaller values of σ_{SOOG} performance in terms of depth becomes very much worse—there is a sharp and sometimes catastrophic collapse in performance. This is presumably a consequence of a limitation in orientational resolving power. There is a minor anomaly in the data for PL at 1000 msec. Here there is an increase in the depth required once σ_{SOOG} is >15 deg, but its magnitude is small compared to the magnitude needed at a σ_{SOOG} of 10 deg.

A simple mathematical description of the data, or model, is suggested by the prominent features of these results. We illustrate it in Fig. 6. The idea is that the visual system has some kind of internal blur over orientation which can be modelled as a Gaussian filter. The Gaussian filter is in the orientational domain and not the spatial domain. This is then convolved with the input representation of the stimulus over orientation—i.e. with the SOOG pdf. This will then demodulate the amplitude of the central dip to a greater or lesser extent, depending on the precise forms of these two functions. We then assume that threshold performance occurs at some criterion depth in the output of the convolution. Thus, for each subject/duration combination we took the results for all the orientational standard deviations (i.e. σ_{SOOG} or σ_U) and convolved the pdfs at threshold d with different widths of the Gaussian filter, σ_F, until we found the most consistent depth (in the sense defined below), in the output. This gives a standard deviation for the model filter and a criterion depth in the output. With these, it is then possible to go back and calculate what the depth at threshold would have been, assuming that the model was operating with those parameters. In Fig. 5 the predictions of the models are shown, along with the orientational standard deviations, σ_F, of the appropriate model filter. This model is essentially the same as that used to understand the detection of orientation modulation (Keeble *et al.*, 1995a, b), except in that case the filter was obtained by a novel use of Fourier analysis in the orientation domain, and did not have an exactly Gaussian form.

Some technical details of the convolution process and model should be mentioned. Because the orientational domain is cyclical, the convolution integral should occur over π radians, taking into account the wrap-around of one or other of the functions of the convolution. We did perform the integral over this range, but in general, because the SOOGs and the optimal filter were always reasonably compact with respect to the π radians domain we assumed that both functions did not wrap around. We tested this approximation by doing the full calculation at selected extreme parameters, and found that it did not affect the model values of σ_F obtained. By "most consistent depth in the output" we mean that we minimize the quantity:

$$\chi^2 = \sum_{i=1}^{N} \frac{(D_{mean} - D_i)^2}{Err_i^2 D_{mean}^2} \quad (3)$$

where N is the number of thresholds obtained for that subject at that duration, D_i is the depth of the dip in the convolution of the pdf at threshold with the filter, D_{mean} is the mean across N of those depths, and Err_i is the difference between the convolved depths for the extrema of the 67% confidence intervals for the depth (i.e. an estimate of the appropriate error weighting for that data point).

This model captures the main features of the data quite well, as can be seen in Fig. 5. The asymptotic behaviour at high σ_{SOOG} and the collapse at low σ_{SOOG} are both reproduced. There appears to be a slight tendency for the model to underestimate the low σ_{SOOG} fall-off in performance in three of the six data sets, possibly indicating that the model is not flexible enough to fit the sharpest collapses in resolution. Nevertheless, for a one-parameter fit to the data it would be unreasonable to expect the model to perform better.

We now turn our attention to the half-height full-width orientational bandwidths ($2.35\sigma_F$) of the model filters emerging from the fitting process. They range from about 10–20 deg, with an across-subject mean of 13.0 deg for the 1000 msec condition, and 17.6 deg for the 105 msec condition. On the one hand, this should be compared with orientational acuities for single lines and sine-wave gratings which are typically < 1 deg (Heeley & Buchanan-Smith, 1990). In a task where a field of Gabor patches with one orientation had to be segmented from a background of patches of a different orientation, Caelli and Moraglia (1985) found thresholds of 3–5 deg. So the mechanism that is at work in producing our results is obviously much coarser than those. On the other hand, when we performed similar experiments involving discrimination between orientationally random textures and orientationally modulated textures (Keeble *et al.*, 1995a, b), we found much broader filter bandwidths of, on average, 34 deg for the 1000 msec condition. What is more, Foster and Ward (1991), have modelled orientational pop-out using orientational filters with a bandwidth of approximately 60 deg. It is clear that our results fall between these extremes. Interestingly, Graham *et al.* (1993), found that their model produced differing estimates of bandwidth for their simple channels (5–20 deg) and their complex channels (more than a factor of 2 higher). The difference in our filter widths between the different exposure durations (a factor of 1.36) could be explicable by cognitive factors, by temporal probability summation, by changes in low-level filter characteristics with time, or by the evolution of grouping processes with time.

In summary, the data are consistent with the view that the magnitude in the dip of the pdf convolved with a relatively broad filter is the determinant of performance in this task. Given that for fixed orientational variance the

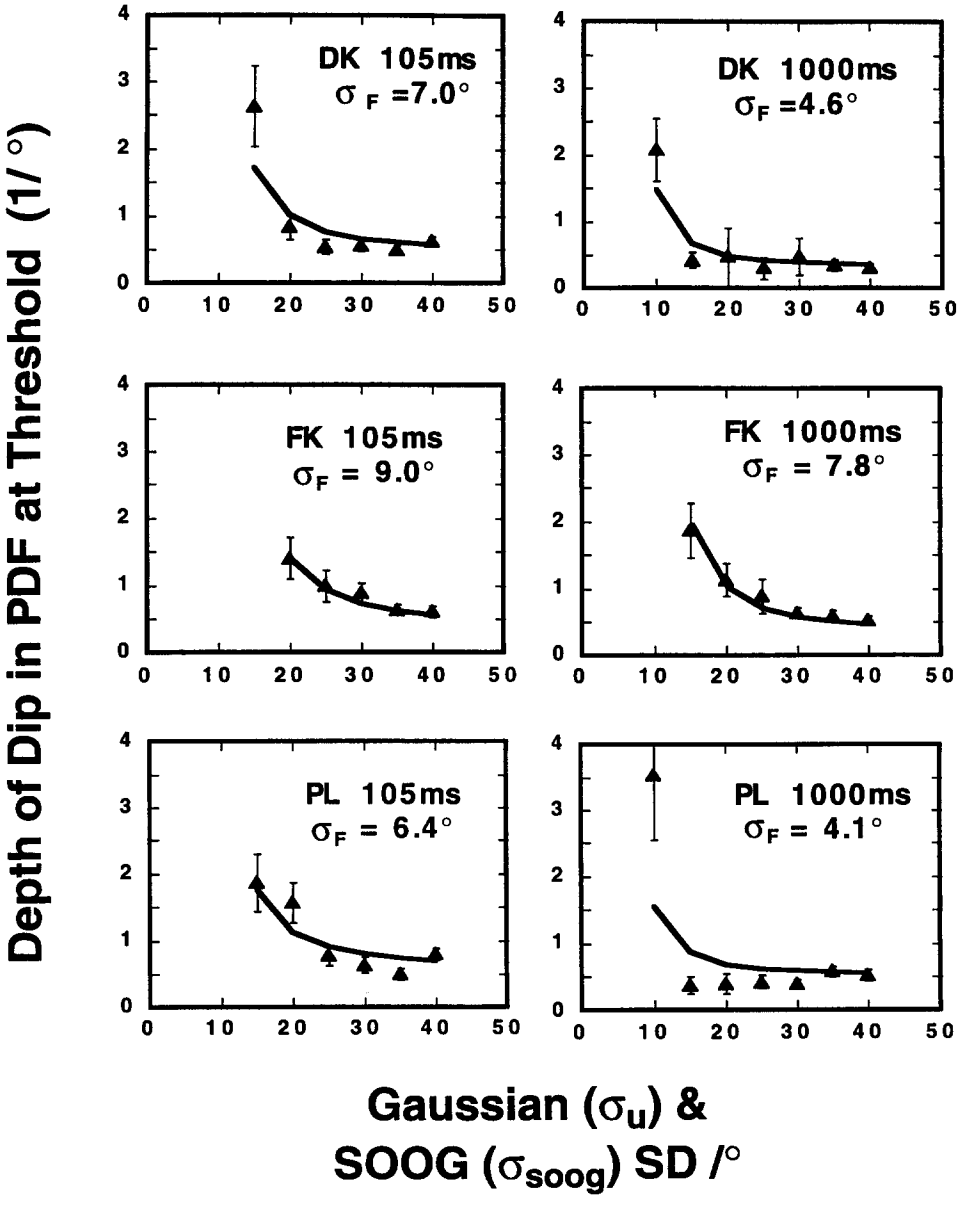

FIGURE 5. Resolution thresholds in Experiment 1 expressed in terms of the depth of the bimodal SOOG, together with model fits (the thick lines), and associated model filter standard deviations (σ_F).

depth of the pdf varies non-linearly with the separation between the peaks, d, and that we found greatly differing values of β for the fits in equation (2), replications of this kind of experiment should use depth as the abscissa variable of the psychometric functions. Attempts to analyse performance in terms of the relative depth of the SOOGs (the dip of the bimodal pdf divided by the peak height of the pdf), rather than the absolute depth did not prove fruitful.

Experiment 2—An absolute resolution limit

Using the experimental paradigm of Experiment 1 it was not easy to find a definitive limit to the orientational texture resolution of the visual system. We can say that when the overall standard deviation becomes < 10–15 deg it becomes impossible to measure thresholds for detecting bimodality, but it is not clear that we have used the optimal bimodal stimulus for eliciting maximal resolution performance. What one should attempt to do is give the observer their best opportunity for performing the discrimination task. In Experiment 2 we do this by just using two discrete orientations as the bimodal pdf. In other words, the pdf is simply two impulse functions. Again the comparison unimodal pdf is a variance-matched Gaussian. The angular distance between the impulse functions is varied, and the percent correct discrimination measured. Clearly, as the difference in orientation between the two angles increases performance will improve. This pdf, and the textures it generates are illustrated in Fig. 3.

The psychometric functions for three observers at two durations are plotted in Fig. 7, along with the Weibull fitted thresholds. For one observer (DK) results were also taken for a duration of 2000 msec. Even for the longest durations a difference of at least 13 deg is required between the discrete orientations for 82% correct

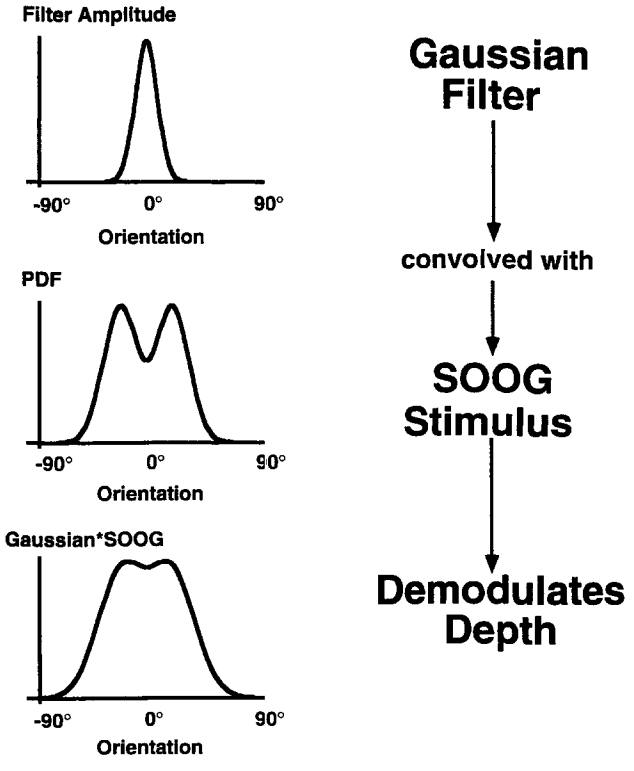

FIGURE 6. Schematic illustration of the orientational filter model.

performance. This is again completely different to the < 1 deg thresholds for single lines and gratings. Testing the model we deployed in Experiment 1 on this data is the logical and straightforward next step. To do this, one takes the model filter Gaussians at the widths given in Fig. 5 and convolves them with two-impulse function pdfs with varying separations. The overall area of the pdf is unity, so each impulse function is normalized to have an area of a half. When the depth in the convolved function is that of the criterion depth for that subject and duration in Experiment 1 then that separation is the threshold prediction. Predictions and actual thresholds are plotted against each other in Fig. 8. They correspond reasonably well, although the threshold separations are, on average about 4 deg larger than predicted. This could be caused by the frequent undershoot in the model fit at low σ_{SOOG} in Fig. 5.

This threshold of 13 deg implies that we are tapping a spatially integrative texture mechanism. Because there are only two orientations in the bimodal texture, and as the orientation of lines can be encoded to an accuracy of better than 1 deg near the fovea (Mäkelä *et al.*, 1993), a mechanism that simply sampled the orientations of a very small number of micropatterns in each array near the fovea with this accuracy would have a much lower threshold than was found.

How "absolute" is this threshold of 13 deg? Or, to restate the question, are there parameters of the stimulus that could be adjusted to improve performance. It seems from DK's results that performance has reached asymptote at 1000 msec, so presumably yet longer durations would make little difference, unless the duration was so long as to allow an intensive serial search of the elements. In pilot experiments for other texture experiments using somewhat similar stimuli we independently manipulated the area of the stimuli and the micropattern density. At the density and area we use in the experiments reported here the performance had reached a plateau. The other plausible determinant of performance is the form of the micropatterns used.

Experiment 3—Influence of the micropattern

We use short line segments as micropatterns because they seem naturalistic to us. For example, images of vegetation generally include many more or less randomly positioned line elements and edges. Lines are broadband in the spatial frequency domain, which has the advantage that the diminution of orientational acuity with eccen-

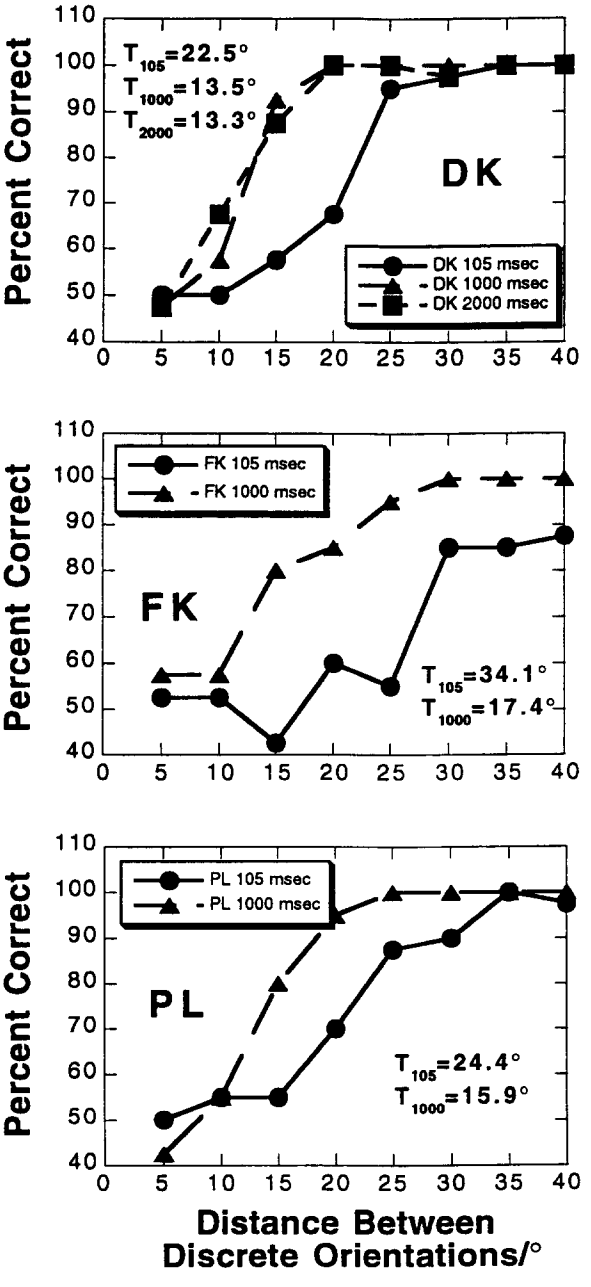

FIGURE 7. Psychometric functions and thresholds for Experiment 2. Performance is plotted as a function of the difference between the two discrete orientations.

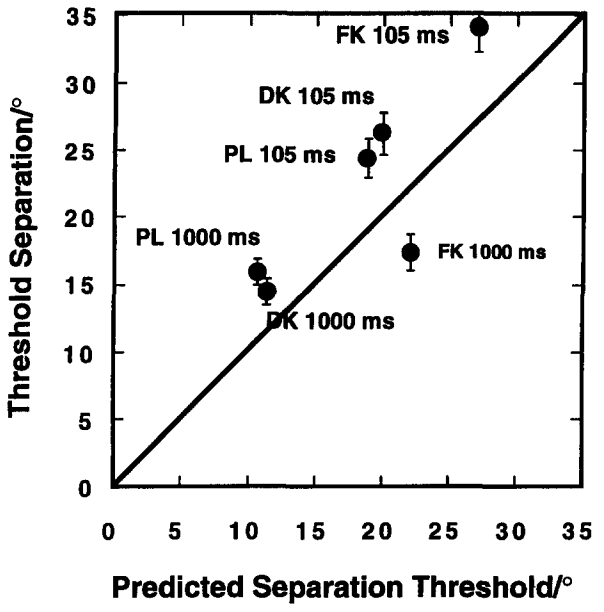

FIGURE 8. Discrimination threshold from Experiment 2 plotted against predicted thresholds from model and the results of Experiment 1.

tricity is less than for narrow-band elements. On the other hand, the early stages of the primate striate cortex are spatial frequency selective, so many workers have argued that spatial vision psychophysics should be conducted using stimuli which are matched to this selectivity. This allows different spatial frequency mechanisms to be independently activated. A popular stimulus for this purpose is the so-called Gabor patch (e.g. Graham, 1989 p. 47), which has a Gaussian profile in one direction, and in the orthogonal direction is a Gabor function—i.e. a Gaussian-windowed sinusoid. This function minimizes, in a certain sense, the joint spread of its position space and frequency space representations. It has also been claimed that it is a reasonable approximation to the receptive fields of striate neurons (Marcelja, 1980). We employ them because: (1) they are spatial frequency narrowband in comparison to line stimuli; and (2) because it is very simple to calculate and to vary their orientational bandwidth. If the spatial frequency content or the orientational bandwidth are important factors in this experiment then we would expect to see differences between the results using line elements and Gabor patches, and also between the results for Gabor patches of differing orientational bandwidths.

In the experiments using Gabor patches, we placed them on a jittered grid as described in the Methods section and shown in Fig. 3(b). This is to allow high contrast values to be used. If random placement was used, lower contrasts would have to be used in order that the luminances of overlapping patches could be added without exceeding the limits of the colour table. In one series of experiments, the peak spatial frequency of the micropatterns was 4.1 c deg^{-1}, the standard deviation of the Gaussian was 8.8 min arc, and thus the orientational half-height full-width bandwidth was 35.8 deg, calculated using the approximation of Graham (Graham, 1989, p. 62). A control condition using line elements on a grid was also used, as depicted in Fig. 3(c). Otherwise, the experimental set-up was identical to Experiment 2—impulse function pdfs with percent correct being measured as a function of angular separation. The results

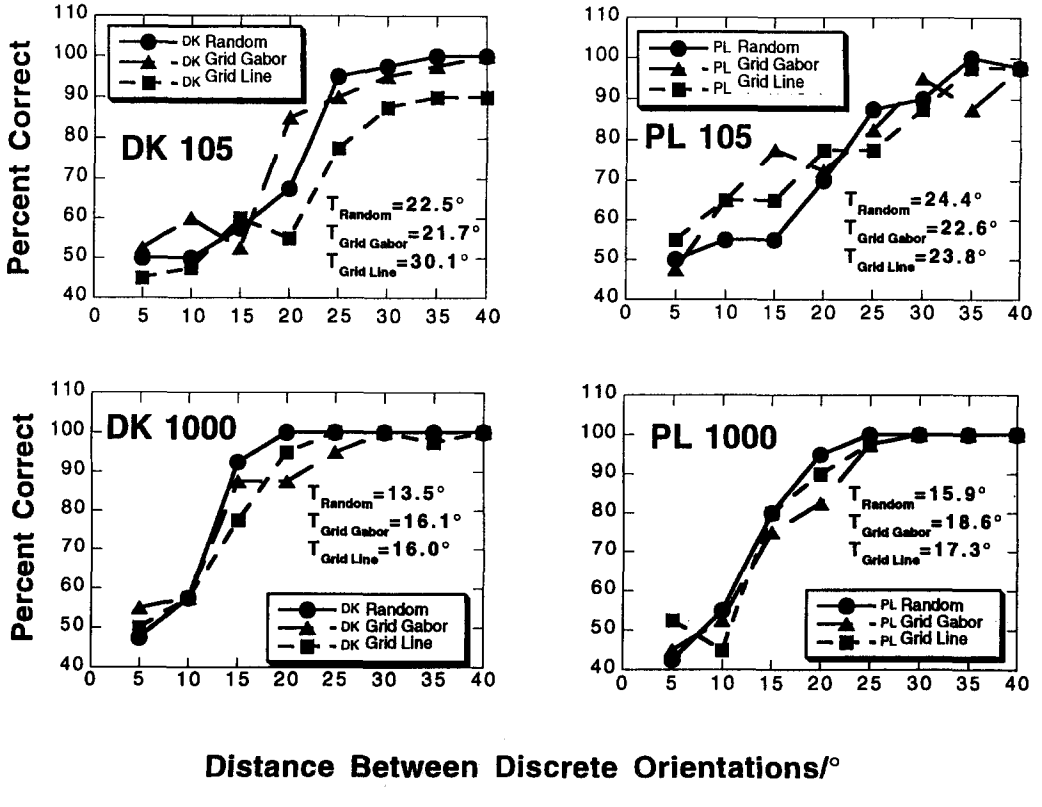

FIGURE 9. Psychometric functions and thresholds for first part of Experiment 3.

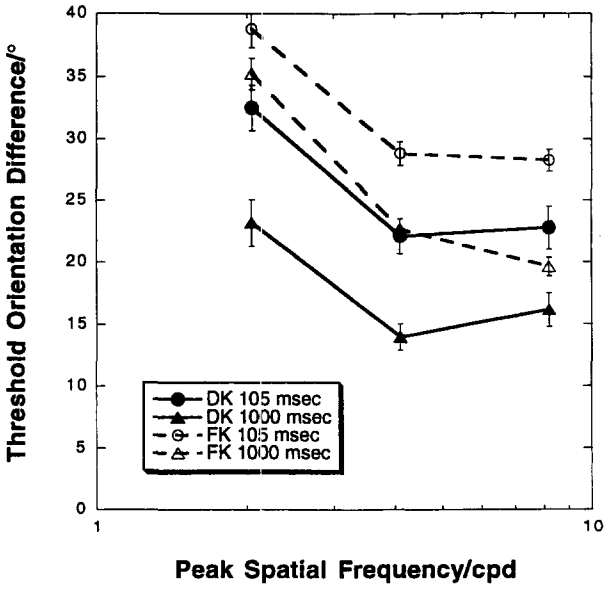

FIGURE 10. Threshold discrimination for Gabor stimuli as a function of peak spatial frequency.

for these experiments can be seen in Fig. 9. In a separate experiment, the orientational bandwidth of the Gabors was varied by changing the peak spatial frequency (2.0, 4.1 and 8.2 c deg^{-1}), whilst keeping the standard deviation of the Gaussian constant at 8.8 min arc. This produced orientational bandwidths of 71.6, 35.8 and 17.9 deg. We show the thresholds for this manipulation in Fig. 10.

There is virtually no difference between the psychometric functions for 35.8 deg bandwidth Gabor patch stimuli and line stimuli shown in Fig. 9. There is a noticeable reduction in performance by subject DK for the grid line condition at 105 msec. As this function seems to asymptote to 90% correct, and as the randomly placed line condition is almost identical to the Gabor condition, presumably this difference is some peculiarity of the grid stimulus. Evidently the form of the spatial frequency representation is not important in this texture task. It is important to stress how different these representations are for lines and Gabor patches. The Fourier transform of an idealized line is the product of two orthogonal sinc functions, with scales inversely related to the dimensions of the line. This will be an extended structure in frequency space which will have its maximum at the origin. By contrast, the Fourier transform of a symmetrical Gabor patch will be a Gaussian blob centred at a point away from the origin of radial distance equal to the frequency of the sinusoidal carrier.

In Fig. 10 the discrimination thresholds for stimuli comprising Gabor patches are plotted as a function of peak spatial frequency. As the size of the Gaussian is constant, orientational bandwidth is inversely proportional to spatial frequency, so the results would look similar if plotted against log orientational bandwidth. Two features are evident here: little change in performance between 4 and 8 c deg^{-1}, and a large decrease at 2 c deg^{-1}. The orientational bandwidth at 2 c deg^{-1} is 71.6 deg, so this decline is not surprising and probably indicates that early stages of the visual system are not being given enough orientational information. In orientational acuity experiments involving arrays of Gabor patches it was found that performance was approximately constant once the bandwidth was smaller than a certain value and that above this value performance collapsed dramatically (Demanins *et al.*, 1996). The bandwidth at which collapse occurred was approximately 60–80 deg. Those results clearly parallel the ones presented here. Those kinds of results have been interpreted as indicating that orientation acuities are not just determined by oriented detector limitations but also by more central factors (Bowne, 1990; Burr & Wijesundra, 1991). Our results can be interpreted similarly.

We mention some of the limitations of this experiment. Varying the spatial frequency of the Gabor micropatterns produces changes in their salience, even at high contrasts, and particularly at more eccentric positions. This is probably the reason for the small increase in threshold for subject DK going from 4 to 8 c deg^{-1}. If the peak spatial frequency was increased still further it seems likely that performance would collapse. As texture stimuli are, of necessity, imaged at various different retinal eccentricities, it seems to us that where possible broad-band stimuli are preferable. We should also mention that obviously the experimental results and modelling reported here can only be averages over different retinal eccentricities, even for the line stimuli, and that there might well be significant variation in performance and model width if a paradigm was employed which allowed the effect of eccentricity to be probed. Because in these experiments the bimodal textures have only two orientations, the patterns of local orientation contrast between the bimodal and the unimodal textures are different. That is, outliers in the Gaussian distribution could conceivably produce local orientation-based pop-out, which would not happen in the two discrete orientation texture. This is a potential alternative mechanism for performing the task. Without performing extensive control experiments it is not possible to completely discount this effect, but it seems from the success of the model in predicting the results in Experiment 2 that the most parsimonious hypothesis is that the effect of local orientation contrast on performance in Experiments 2 and 3 is small.

In summary, the result of Experiment 3 is that performance is not greatly affected by the kinds of micropattern used, except when the orientational bandwidth is made very large.

DISCUSSION

We commence by summarizing the principal findings of these experiments:

- Orientational bimodal texture features must be separated by at least 13 deg in order to be discriminated from unimodal textures with the same variance.

- For high variance, large separation textures, performance is governed by the size of the central dip in the bimodal pdf, whilst at lower separations and variances larger dips are required.

- Convolution of the bimodal pdf with a Gaussian filter together with the assumption that performance depends on the dip in the output models observer performance well, giving filter full-width bandwidths of about 10–20 deg.

- The form of the micropatterns employed is not an important factor in determining performance.

Some related measures of resolution performance can be obtained from our previous work (Keeble *et al.*, 1995a, b) in which orientationally modulated textures had to be discriminated from orientationally random textures. If the orientationally modulated texture was generated from a sinusoidal pdf, then threshold performance for maximum modulation could, on average, only be achieved if the separation between peaks was more than about 65 deg. This is the analogue of the conventional resolution limit in the luminance domain (Bennett & Rabbetts, 1984). If instead equispaced discrete orientations were used, threshold was reached on average at a spacing of 32.2 deg The difference between these limits again demonstrates that the difference in angles between the peaks is not the key variable which determines performance. In those experiments the results were adequately modelled with an orientational filter of average width 34 deg for the 1000 msec condition.

The modelling of the results in this paper gives a mean filter width for the 1000 msec condition of 13.0 deg. What are the possible reasons for this difference of more than a factor of 2.5? It seems clear that a simple linear model with one filter width and performance based on the amplitude of the output modulation cannot simultaneously explain the two sets of results. The essential difference between our previous results and those reported here is that in these experiments the orientational content is always massively above threshold (that is, there is always a large anisotropy), whereas, before, the orientational content was at threshold for discrimination from an orientationally random texture. This means that in the resolution task subjects know in what general angular region the difference between the pdfs lies, whereas in the detection task subjects must search over 180 deg for the modulation from random. If this attentional kind of explanation is correct, then interpretation of the orientational filter as a low-level process would have to be treated with some caution.

An alternative explanation for the differences in parameter values observed could be linked to the interpretation of the filter functions obtained in the two sets of experiments. We will argue later in this section that the limitation to resolution is not a first-stage filter limitation, but represents higher processes. These could either be dedicated second-stage filters or grouping operations. That is, lines with similar orientations would group together. There is a plethora of evidence for the existence of lateral connections between neurons tuned to similar orientations (e.g. Ramoa *et al.*, 1986; Ts'o *et al.*, 1986; Ts'o & Gilbert, 1988; Gilbert & Wiesel, 1989), whilst the concept of grouping by similar orientation is well-known to Gestalt psychology (Wertheimer, 1938). Excitation between similarly tuned neurons would act to "sharpen up" the cortical representation of orientationally bimodal textures, provided the difference between peaks is sufficiently large, otherwise, grouping would tend to merge the two peaks. In this way the linear model filters which we found would be regarded as epiphenomena of grouping processes, with the width of the filter reflecting the strength of the grouping. If this is the underlying mechanism, then it is interesting that a complex, intrinsically non-linear process can be modelled in a simple linear way. It might be that for textures with supra-threshold orientational content grouping takes place more effectively or quickly than for textures which are at modulation threshold.

In fact, the coarseness of texture resolution that we find here is not surprising from an ecological point of view. It is probably not very important for an organism to have a precise knowledge of the shape of the orientation distribution of a texture. What would seem to be useful is the direction, strength and variance about the mean of the distribution [see discussion section in Keeble *et al.* (1995a)]. It is worth reiterating that orientation acuities for lines and gratings can be better than 1 deg, and so would not predict our results here, particularly those in Experiments 2 and 3 where the bimodal textures comprise only two orientations. It would be interesting to know whether orientational acuities (that is, ability to assess the overall orientation of a textured region) for textured stimuli are nearer to the < 1 deg limit for single patterns than the 13 deg resolution limit reported here. Our and other observations suggest that this is so, (Keeble *et al.*, in prep.; Dakin, 1994).

In Experiment 3 it was evident that the nature of the micropattern was not an important factor in performance: lines and Gabor patches with less than a certain orientational bandwidth produced similar results. This indicates that performance is probably not governed by the primary responses of first-stage oriented filters, or in other words, V1 orientationally sensitive neurons. This parallels the findings of Bowne (1990) that central noise (i.e. noise at a later stage of the system), is important in spatial frequency, orientation and temporal frequency discrimination. Most models of human texture perception have been couched in terms of second-stage filters which integrate information over some region (e.g. Landy & Bergen, 1991; Malik & Perona, 1990; Graham *et al.*, 1992). An alternative could be the grouping processes we mentioned earlier. Whichever is correct, the non-dependence of human performance on the input spatial frequency characteristics strongly suggests that the resolution limitation occurs at the later stage.

We have utilized the orientational pdf as a means of describing how the strength of the orientational signal in

the texture varies with orientation. The other obvious descriptor of orientational power is the Fourier transform of the texture as a function of orientation. However, the relative invariance of performance with micropattern type and hence with amplitude spectrum, indicates that the pdf description is superior, at least for these experiments.

In conclusion, we finish by emphasizing that the unanalysed orientational pdf of a texture cannot contain all the secrets of texture perception: it will only reveal certain aspects of it. In particular, the pdf does not make explicit the local orientational contrast which will occur in line textures. It has been known for many years that such contrast is very salient, and can be the basis for pop-out. In this paper we have discussed bimodal textures that appear unimodal (because of resolution limitations), but, conversely, there may well be unimodal textures which appear bimodal because of local orientation contrast effects. These will be the subject of future investigations.

REFERENCES

Beck, J. (1966). Effect of orientation and of shape similarity on perceptual grouping. *Perception and Psychophysics, 1*, 300–302.

Beck, J., Prazdny, K. & Rosenfeld, A. (1983). A theory of textural segmentation. In *Human and machine vision* (pp. 1–38). London: Academic Press.

Bennett, A. G. & Rabbetts, R. B. (1984). *Clinical visual optics*. London: Butterworths.

Bowne, S. F. (1990). Contrast discrimination cannot explain spatial frequency, orientation or temporal frequency discrimination. *Vision Research, 3*, 449–461.

Burr, D. C. & Wijesundra, S. A. (1991). Orientation discrimination depends on spatial frequency. *Vision Research, 7/8*, 1449–1452.

Caelli, T. & 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, 197*, 551–566.

Dakin, S. C. (1994). The visual representation of texture. Ph.D. Thesis, University of Stirling.

Demanins, R., Hess, R. F. & Keeble, D. R. T. (1996). The nature of the orientation deficit in strabismic amblyopia. *Investigative Ophthalmology and Visual Science, 37 Supplement*, S280.

Foster, D. H. & Ward, P. A. (1991). Horizontal-vertical filters in early vision predict anomalous line-orientation identification frequencies. *Proceedings of the Royal Society of London B, 243*, 83–86.

Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.

Gilbert, C. D. & Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and cortico-cortical connections in cat visual cortex. *Journal of Neuroscience, 9*, 2432–2442.

Graham, N. (1989). *Visual pattern analyzers*. Oxford: Oxford University Press.

Graham, N., Beck, B. & Sutter, A. (1992). Nonlinear processes in spatial-frequency channel models of perceived texture segregation: Effects of sign and amount of contrast. *Vision Research, 32*, 719–743.

Graham, N., Sutter, A. & Venkatesan, C. (1993). Spatial frequency and orientation-selectivity of simple and complex channels in region segregation. *Vision Research, 33*, 1893–1911.

Heeley, D. W. & Buchanan-Smith, H. M. (1990). Recognition of stimulus orientation. *Vision Research, 30*, 1429–1437.

Hubel, D. H. & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology (London), 148*, 574–591.

Julesz, B. (1962). Visual pattern discrimination. *IRE Transactions on Information Theory, 8*, 84–92.

Keeble, D. R. T., Kingdom, F. A. A., Moulden, B. & Morgan, M. J. (1995). Detection of orientationally multimodal textures. *Vision Research, 35*, 1991–2005.

Keeble, D. R. T., Kingdom, F. A. A. & Morgan, M. J. (1995). The orientational resolution of texture perception. *Investigative Ophthalmology and Visual Science (Suppl.), 36*, S229.

Keeble, D. R. T., Kingdom, F. A. A. & Morgan, M. J. (1988). The perceived orientation of line textures. In preparation.

Landy, M. S. & Bergen, J. R. (1991). Texture segregation and orientation gradient. *Vision Research, 31*, 679–691.

Levi, D. M. & Klein, S. A. (1990). Equivalent blur in spatial vision. *Vision Research, 30*, 1971–1993.

Mäkelä, P., Whitaker, D. & Rovamo, J. (1993). Modelling of orientation discrimination across the visual field. *Vision Research, 33*, 723–730.

Malik, J. & Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America A, 7*, 923–932.

Marcelja, S. (1980). Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America, 70*, 1297–1300.

Ramoa, A. S., Shadlen, M., Skottun, B. C. & Freeman, R. D. (1986). A comparison of inhibition in orientation and spatial frequency selectivity of cat visual cortex. *Nature, 321*, 237–239.

Ts'o, D. Y. & Gilbert, C. D. (1988). The organization of chromatic and spatial interactions in the primate striate cortex. *Journal of Neuroscience, 8*, 1712–1727.

Ts'o, D. Y., Gilbert, C. D. & Wiesel, T. N. (1986). Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. *Journal of Neuroscience, 6*, 1160–1170.

Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics, 18*, 292–297.

Wertheimer, M. (1938). *Laws of organization in perceptual forms*. London: Harcourt (Brace and Jovanovich).

Acknowledgements—This work was supported by the Image Interpretation Initiative of SERC (U.K.), Grant Reference GR/F 37124; and NSERC (Canada), Grant Reference OGP 0121713. Some of the results were presented at the Association for Research in Vision and Ophthalmology, 15 May 1995, and have been published in abstract form (Keeble *et al.*, 1995b). We wish to thank Rita Demanins and R. Eric Fredericksen for their assistance.