



The Computation of Orientation Statistics from Visual Texture

S. C. DAKIN,*† R. J. WATT*

Received 7 October 1996; in revised form 6 March 1997

This paper examines how observers estimate the overall orientation of spatially disorganised textures containing variable orientation. Experiments used asymmetrical distributions of orientations to separate the predictions from different models of average orientation estimation. Stimuli were composed of two spatially intermingled sets of oriented patches, each set having Gaussian distributed element orientation. The threshold separation of the means of the two sets was determined for a variety of tasks. Discrimination of these textures from a reference composed of two sets with the same mean orientation was well predicted by discrimination of orientation variability. A single interval judgement of which set contained more elements required a greater separation of the set orientations and suggested that the sets must be resolved in the orientation domain for independent representation of their properties. That resolution is required to perform this task further suggests that orientational skew is not coded. Threshold offsets for judgement of average orientation were re-expressed as shifts of four candidate features for coding the central tendency of texel orientations. Comparison with similar thresholds for single distributions of orientations indicated that average orientation is assigned to the centroid of a set of orientation measures. © 1997 Elsevier Science Ltd

Texture Orientation Scale Centroid Resolution

INTRODUCTION

Most psychophysical studies of texture perception, largely influenced by the work of Julesz (e.g. Julesz, 1981), have concentrated on the extraction of texture boundaries. That work presupposes that texture is primarily useful as a supplement to contrast information for the detection of surface discontinuities. Experimentally, boundaries are usually defined as rapid differences between attributes, such as orientation, of adjacent texture elements (texels). Such an approach assumes that, in general, textural cues will be available from local differences between texels. Usually, however, attribute information derived from natural images is variable and must require spatial combination to be useful.

The extraction of global texture attributes—and in particular, orientation—is known to be useful in establishing a number of different surface properties. Kass and Witkin (1985) used such statistics to estimate the formative processes a texture had undergone. Similarly estimates of local surface shape require integration and

calculation of (usually second) moments of local orientation statistics (e.g. Blake & Marinos, 1990; Witkin, 1981). Additionally, segmentation might be achieved not only through local differences in attributes but also in differences between integrated attribute statistics (e.g. difference in local mean orientation: Vorhees & Poggio, 1987). Little is known of human perceptual performance where texture statistics are explicitly varied.

Marr (1982) observed that, in order to understand a complex process such as texture perception, the formation of intermediate representations is required. What are the primitives for representing texture orientation statistics? Recent observations suggest an analogy between the visual processing of luminance features and textural boundaries. It has been shown that texture segregation is sensitive to the rate of change of texel orientation across space (the structure gradient) at field boundaries (Landy & Bergen, 1991; Nothdurft, 1985a, b, 1991; Wolfson & Landy, 1995). Landy & Bergen (1991) used spatially band-pass textures to show that this effect is independent of the local density of texels. In addition, Sagi (1990) found a non-monotonic relationship between performance on a visual search for a vertical target and the number of (horizontal) distractors. Data were well explained using “hyper-filters”, which integrate local orientation measures over a restricted area. Input is

*Department of Psychology, University of Stirling, Stirling, Scotland FK9 4LA, U.K.

†To whom all correspondence should be addressed. at present address: McGill Vision Research Unit, Department of Ophthalmology, 687 Pine Avenue West, H4-14, Montreal, Quebec, Canada H3A 1A1 [Fax +1-514 843 1691; Email scdakin@astra.vision.mcgill.ca].

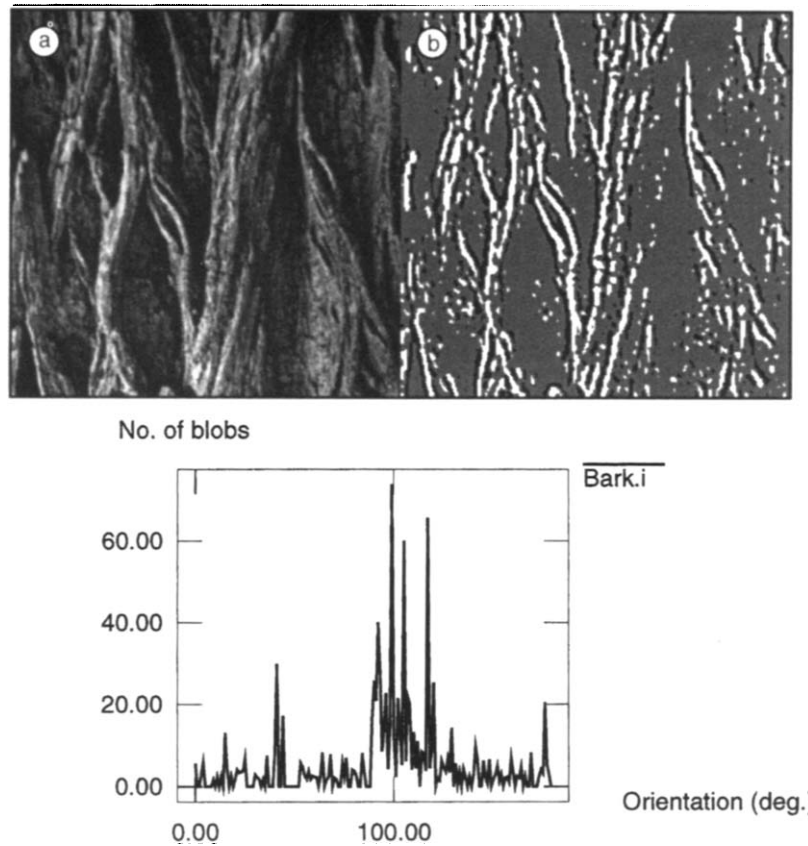


FIGURE 1. (a) Image of tree bark, a strongly oriented texture. (b) A set of possible texture elements. The "blobs" shown are derived by first convolving the image with an isotropic, Laplacian-of-Gaussian filter and then retaining only high energy regions (specifically, those pixels whose values are at least one standard deviation greater or less than the mean output value). (c) Orientation histogram of blobs derived from a symbolic description of (b) (Watt, 1991).

independent of spatial frequency, and filters indicate significant changes of these measures in space (Fogel & Sagi, 1989; Sagi, 1990).

These studies suggest that performance on visual search and texture segregation tasks, using orientation cues, may be well explained by mechanisms which extract differences between spatially integrated local orientation measures. Together they strongly implicate a further stage of combination and processing of orientation information beyond the accepted representation at V1 (Hubel & Wiesel, 1967).

Consider Fig. 1 which shows an oriented texture, along with a set of candidate texture elements [Fig. 1(b); derived using the thresholded output of a Laplacian-of-Gaussian filter] and a histogram of the orientation of these blobs [Fig. 1(c)]. One is capable of correctly estimating that the average orientation of the original texture is approximately vertical, even though the histogram shows that there is a great deal of variability in blob orientation. If one were interested in an orientation representative of the overall direction of surface creases (for example, to position the cutting edge of a tool on the surface) one would have to use some form of combination of measures to calculate it. Any individual texel would be too unreliable. The question that this paper addresses is this: which orientation statistics does the human visual system compute and

particularly, how does it represent the central tendency of orientations?

Observation of how a system's performance deteriorates with the addition of noise is a powerful approach for understanding that system (see, e.g. Barlow, 1980). By employing textures composed of separable texels one may examine the effect of noise explicitly added to the orientation of components. This approach to understanding the perception of average orientation of textures has already been applied elsewhere (Watt, 1991). When the orientation of texels varies around some mean value, judgement of average orientation is limited only by the variance of texel orientations and by some constant internal error of representation. By further supposing that orientations are averaged, an estimate of the number of texels employed can be made. Such a model accounts for the data presented and suggests that the mean of a number of orientation measures was available to subjects. Thus, Watt (1991) offers preliminary evidence that integration of texel orientation occurs in the judgement of average orientation of a texture field. Furthermore, this work suggests that the arithmetic mean of texel orientations predicts thresholds well.

The remainder of the paper explicitly compares alternative strategies for coding average orientation. Is averaging of orientations required at all or could an alternative feature, such as the most numerous orienta-

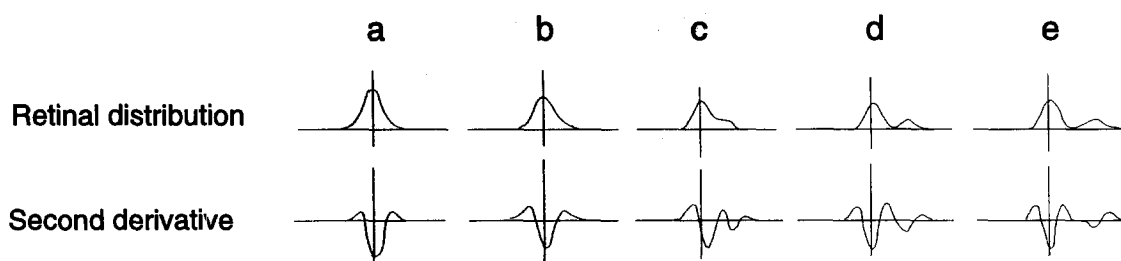


FIGURE 2. The upper row shows the retinal light distributions of typical stimuli from Watt & Morgan (1983), the lower row shows the second derivatives of these distributions. (a) Shows the distributions from two superimposed bright bars with luminance ratios of 2:1. (b–e) The effect of increasing the separation of the bars.

tion, suffice? In addition, the representation of multiple oriented sets is investigated. An understanding of these issues will give insight into the computation of texture statistics.

In the experiments reported in this paper all stimuli were spatially unstructured single patches of oriented texture. The experimental procedures and models employed explore an analogy between the processing of luminance features and texture. Assume that an estimate of the underlying orientation probability density function (p.d.f.) is formed through averaging of local texel attributes, i.e., an orientation histogram is constructed. It is hypothesized that the extraction of statistical moments (e.g. central tendency) of this representation will occur in a similar way to the extraction of moments from retinal luminance distributions. In the case of the central tendency of a retinal distribution two stages are involved in most models: the integration of luminance values through a blurring function and the accentuation of some moment of the resulting distribution (e.g. by taking the second derivative). It is proposed that in forming an estimate of the central tendency of an orientation distribution, similar processes may be used.

METHODS

Rationale

Tasks exploring the location of spatio-luminance distributions (Watt & Morgan, 1983), were adapted to look at the coding of orientation distributions.

Watt and Morgan (1983) employed tasks using spatially combined bright bars to compare the predictive power of four visual location models. By varying the relative brightness of component bars, the asymmetry of the compound feature presented could be varied. Using these stimuli the threshold offset of the mean positions of the bars was determined for vernier acuity and resolution tasks. Figure 2 shows the retinal light distributions, and their second derivatives, of typical stimuli from their experiment. The distributions from two superimposed bars, with luminance ratios of 2:1, are shown in Fig. 2(a). Figure 2(b–e) shows the effect of increasing the separation of the two bars. At the resolution point [Fig. 2(c)] the number of zero-crossings in the second

derivative increases from two to four. A model based on the extraction of zero-crossings in the second derivative best accounted for the subjects' accuracy at estimating the position of these asymmetrical luminance distributions. A centroid model also performed well and has been implicated elsewhere in deriving the location of dot clouds (Whitaker & Walker, 1988) and flanked bars (Badcock & Westheimer, 1985). Peak and threshold edge models produced poor fits to the data.

The approach used in Watt and Morgan's (1983) paper was used as a framework for experimental method and the models they examined were adapted to process texel orientation distributions. Rather than using bars defined by two distributions of luminance added in space, the orientation of texels was determined by two probabilistic distributions. Textures will have orientation distributions similar to the upper row of Fig. 2. By shifting the relative mean orientations of the two distributions an asymmetry can be constructed, analogous to the spatio-luminance case.

The textures that were used consisted of two intermingled sets of spatially band-pass, oriented texels (see Fig. 3 for examples). The orientations of elements were randomly drawn from gaussian distributions (SD of 6 deg) and the relative number of texels in each set was systematically varied. Using an adaptive psychophysical procedure the threshold offset of the mean orientations of the two sets was determined for three tasks. They were:

- A resolution judgement: subjects identified the patch that was composed of sets with different mean orientations.
- A judgement of which set within one texture patch was more numerous.
- A comparison of the average orientation of the patch with vertical.

Several models were compared to determine the cues used to make the judgements.

Performance for judging the average orientation of asymmetrical distributions (Experiments 3 and 4) was used to separate predictions from the models, as was the case in Watt and Morgan (1983).

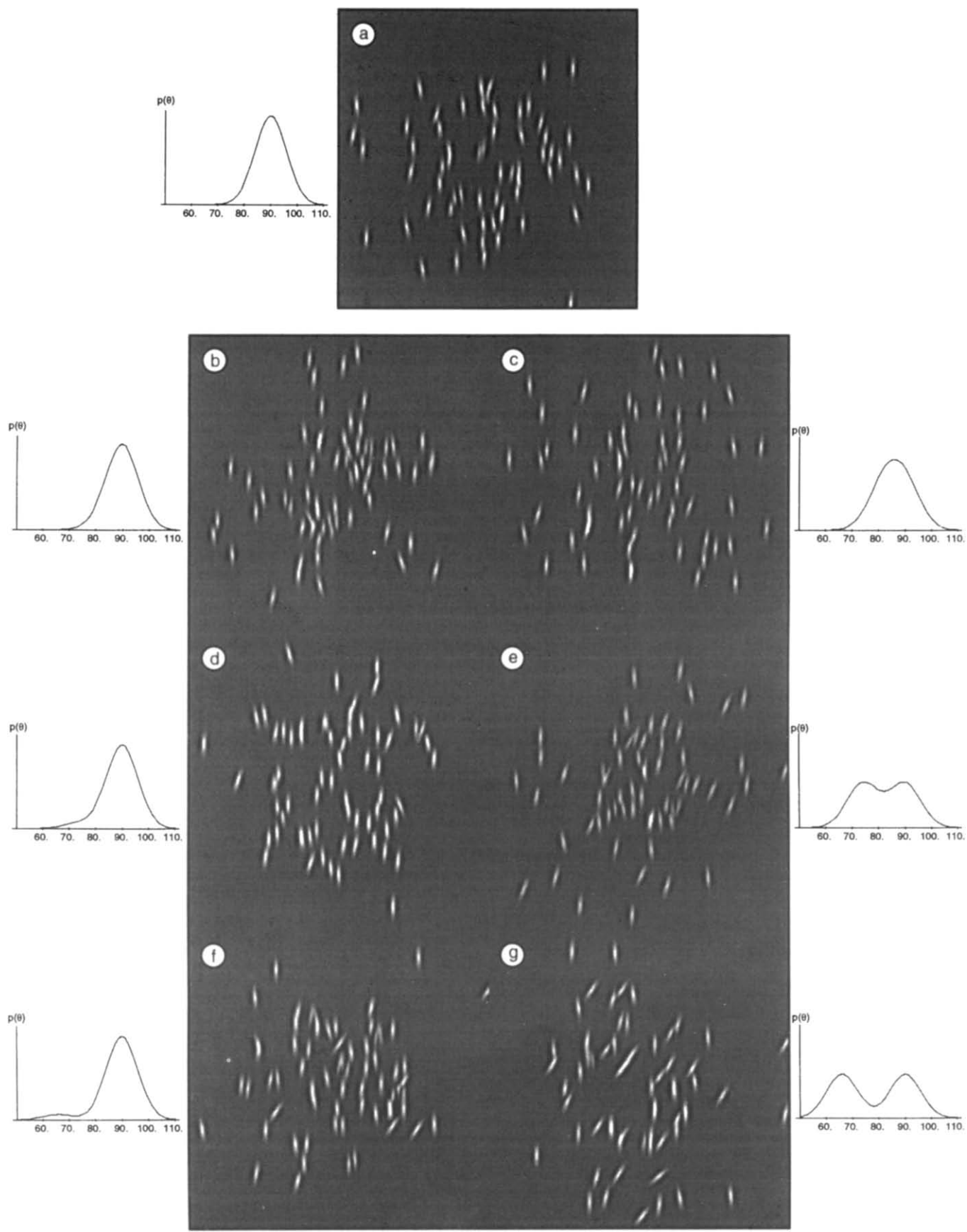


FIGURE 3. Example stimuli from Experiments 1–4 (contrast-enhanced for reproduction). All textures contain 64 elements and the orientation of the more numerous set is 90 deg. (a) shows the reference stimulus, containing two superimposed texel sets (i.e., the cue $\delta\theta = 0$ deg). (b, d, f) Two intermingled textures: each set has the same number of elements ($n_1 = n_2$). The relative offset of the two sets, $\delta\theta$, is (b) 0 deg, (d) 8 deg, (f) 16 deg. (c, e, g) As above, but for a texture where one set has 16-times as many elements as the other ($n_1 = 16n_2$). Relative offsets are (c) 0 deg, (e) 8 deg, (g) 16 deg. Graphs show the p.d.f.s for the adjacent texture.

Subjects

The first author and five observers naïve to the purposes of the experiments served as subjects. Three (SCD, PMC and DFS) were experienced psychophysical observers, while the others had limited experience. DFS and SCD were corrected-to-normal myopes and DFS also had a slight (less than 0.5 D) corrected astigmatism.

Apparatus

All stimuli were presented on a Formac ProNiton 80.21 colour monitor with a frame refresh rate of 80 Hz, and gamma corrected. The screen was viewed binocularly with natural pupils at a distance of 49 cm, and had a mean luminance of 26 cd/m². Subjects fixated the centre of the display, indicated using a pre-stimulus fixation marker.

Stimuli

In order to restrict orientation information to an individual spatial scale, textures with relatively narrow-band spatial frequency characteristics were used. Textures were composed of small patches with luminance modulated by a Difference-of-Gaussian (DoG) function. Many of these DoG elements were randomly distributed throughout the image and added together. They appeared on a mid-grey background. The DoGs making up the textures had a luminance profile given by:

$$L(x, y) = L_0[1 + W(x, y)]$$

where:

$$W(x_t, y_t) = A \left[e^{-x_t^2/2\sigma^2} - \left(\frac{1}{2.23} e^{-x_t^2/2(2.23\sigma^2)} \right) \right] e^{-y_t^2/2(3.2\sigma^2)}$$

$\sigma = 3.48$ min arc and A is a scaling constant. x_t and y_t are the translated and rotated coordinates:

$$x_t = (x - \mu_x)\cos \phi + (y - \mu_y)\sin \phi$$

$$y_t = (y - \mu_y)\cos \phi + (x - \mu_x)\sin \phi$$

μ_x , μ_y represent the translation and ϕ the rotation of the function. This is an adapted sensitivity profile given by Wilson and co-workers (Phillips & Wilson, 1983; Wilson & Gelb, 1984) for the 4 c/deg channel. The DoGs were clipped at ± 3.0 standard deviations from their centres. Overlapping patches were numerically added. The whole image was stored with floating point precision before being normalized to a range appropriate to the display hardware for presentation. Differing degrees of overlap between texels effectively randomized the contrast range from trial to trial.

The texture patch contained a Gaussian random spatial distribution of texels ($\sigma = 50$ pixels), centred on the middle of the display. The texture fitted into a central window of 256 pixels square (10.0 deg square).

The orientation p.d.f.s used to generate the orientation of DoGs were typically composed of the sum of two Gaussians with differing offsets and relative heights, but with consistent standard deviations of 6 deg. Two effects

determined this choice. Firstly, aliasing of the probability distribution due to pregeneration of DoGs in 1 deg steps becomes problematic at very small standard deviations. Secondly, a small number of elements poorly characterize a Gaussian distribution at very large standard deviations.

Procedure

A Macintosh IIfx microcomputer generated and presented the stimuli, and recorded subjects' responses. The subjects' task was always a binary decision involving one (Experiments 2, 3 and 4) or two (Experiment 1) intervals. Unless stated otherwise, textures were presented for 100 msec. For the two interval experiments, the first texture was presented in the centre of the display, followed by a 750 msec delay, followed by the second texture. In the single interval experiments the texture was presented, followed by a 750 msec delay. Subjects indicated their decisions by depressing one of two keys on the computer keyboard. APE, an adaptive method of constant stimuli (Watt & Andrews, 1981), was used to sample a representative range of relative numerosities of the two sets.* Probit analysis was applied to the response data to estimate the standard deviation of the psychometric function. Three runs of 64 trials each were undertaken for all conditions, unless stated otherwise. The data points presented are the arithmetic mean of thresholds from the three runs and error bars are an estimate of the standard error of these data.

EXPERIMENT 1: SEPARATING MIXED TEXEL SETS

Stimuli were composed of two intermingled sets of DoG patches (see Fig. 3 for examples) containing n_1 and n_2 elements, respectively and with the same orientation standard deviation ($\sigma = 6$ deg). The p.d.f. to produce this distribution of orientations was the sum of two gaussians with means of θ and $\theta + \delta\theta$. The orientation p.d.f. for the stimulus is then defined as:

$$p(\theta) = \frac{1}{\sigma\sqrt{2\pi}} \left[p_1 \exp \left[-\frac{(\theta - \theta)^2}{2\sigma^2} \right] + (1 - p_1) \exp \left[-\frac{(\theta - \theta - \delta\theta)^2}{2\sigma^2} \right] \right] \quad (1)$$

Where $p_1 = n_1/(n_1 + n_2)$. In the reference stimulus the means of the two distributions were identical ($\delta\theta = 0$ deg). In the comparison stimulus a non-zero increment was added to one of the means. The threshold $\delta\theta$ was determined for the task of reporting which of the fields had orientations drawn from distinct distributions.

*The version of APE used differed slightly from the procedure described in this reference. The fitting of the response distribution with the cumulative normal error model was performed every trial, and not after each block of trials, as described in the Watt and Andrews (1981) paper. This addition only serves to increase the efficiency of the algorithm.

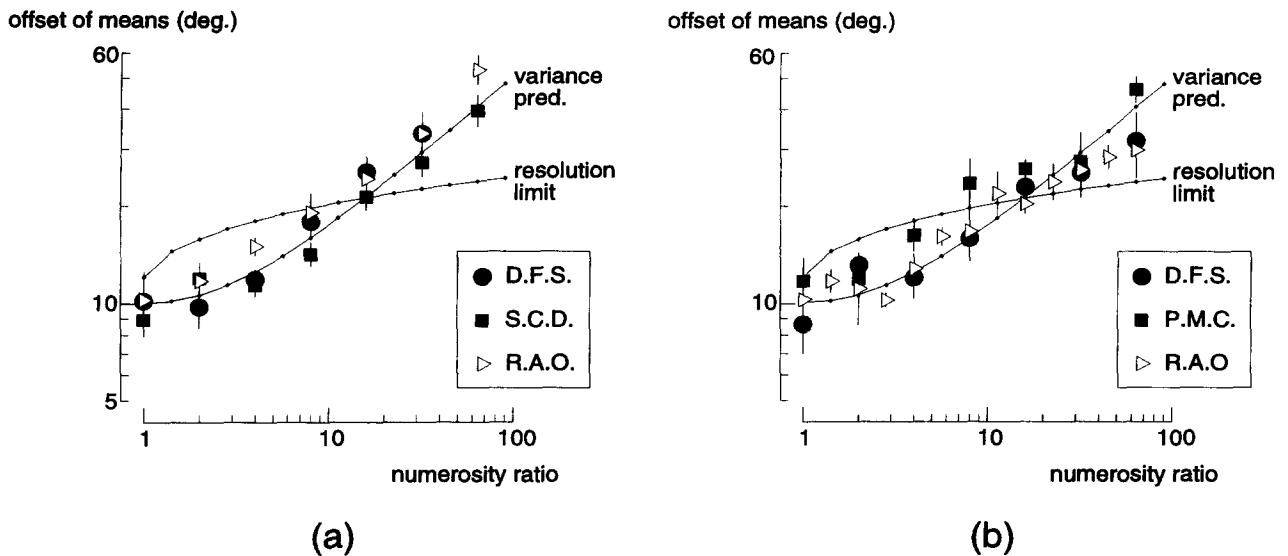


FIGURE 4. Results for three subjects from Experiment 1 with a total of (a) 64; and (b) 256 texels. The abscissa represents the relative number of texels in the two sets (n_1/n_2); the ordinate represents the threshold offset of means in degrees. Predictions from a judgement of the degree of orientation variability and the theoretical resolution limit of the distributions (as defined in the text) are also shown.

Specifically, subjects were asked to discriminate which texture was composed of two sets of elements, with different overall orientations, from the texture which only contained one set.

In order to prevent subjects from using mean orientation cues θ was randomized from trial to trial. The total number of DoG patches, $n_1 + n_2$, remained the same within a condition. Thresholds (as defined above) were determined at relative numerosities ranging from 1:1 to 1:64 evenly spaced on log axes. All subjects were run with 64 and 256 total DoGs.

Results are shown in Fig. 4. Thresholds are lowest with equal number of elements drawn from both sets and increase with relative difference in number. There is little difference between the results from the conditions with 64 and 256 total elements.

Figure 4 also shows the theoretical resolution limit for the stimuli employed in Experiment 1. In the spatio-luminance case, Watt and Morgan (1983) define this as the point at which the number of stationary points in the first derivative of the luminance profile changes from one to three. If it is assumed that the observed distribution of element orientations will approach the probability density functions used to generate them, then the orientation distribution of the patterns may be treated analogously to the spatial case and resolution limits calculated. For low ratios of texel number ($< 1:8$) all subjects perform below this resolution limit for the two distributions. This is in agreement with the findings of Watt and Morgan (1983). They suggested that because performance was not limited by resolution, the task amounted to discrimination of the *width* of luminance distributions. Similarly, the conclusion here is that subjects used a measure of the orientational variability in the texture patch to perform the task at offsets below

the theoretical resolution limit. In other words, this task does not require the detection of bimodality but can be performed using a simple population statistic: orientation variability.

To investigate this more thoroughly, the data were fitted with predictions derived from an orientation variability discrimination task. In this task subjects reported which of two patches, each composed of a single set of DoG patches, had the greater orientation standard deviation. The reference set had orientations drawn from a distribution with $\sigma = 6$ deg, and the comparison set had a larger standard deviation. The threshold increase in standard deviation was determined for discrimination. Textures had randomized mean orientation: other experimental details were identical to those described in the Methods section.

The threshold difference in standard deviation of orientation ($\delta\sigma$) was determined for each subject. The average threshold across subjects was $\delta\sigma = 2.0$ deg. The offset of means producing this rise in the standard deviation of the combined distribution was calculated for each of the numerosity ratios from Experiment 2. The mean (θ_c) and variance (σ_c^2) of two combined distributions is given by:

$$\theta_c = \frac{n_1\theta_1 + n_2\theta_2}{N} \quad (2)$$

$$\sigma_c^2 = \frac{n_1\sigma_1^2 + n_2\sigma_2^2 + n_1(\theta_1 - \theta_c)^2 + n_2(\theta_2 - \theta_c)^2}{N} \quad (3)$$

where n_1 and n_2 are the number of elements in two sets, and N is the total number of elements ($n_1 + n_2$). Sets have means of θ_1 and θ_2 and standard deviations of σ_1 and σ_2 . Given that the standard deviations of sets are equal,

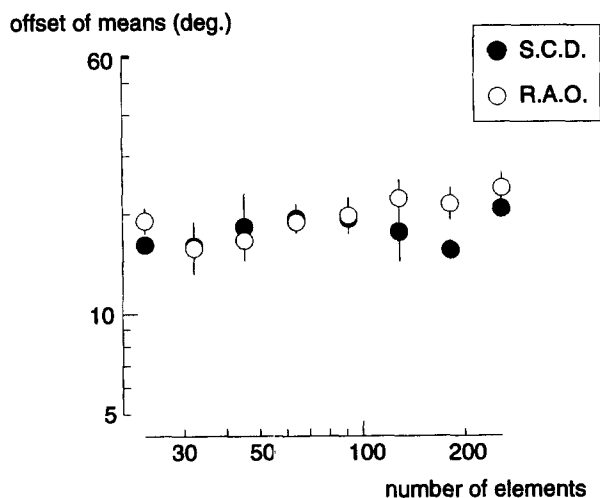


FIGURE 5. The effect of the total number of texels on the task used in Experiment 2. The abscissa represents the total number of texels; the ordinate represents the threshold separation of the set means (in degrees) for discrimination of textures composed of two texel sets from those composed of one.

$\sigma = \sigma_1 = \sigma_2$ is substituted in Eq. (3), and combined with Eq. (2) to give:

$$|\theta_1 - \theta_2| = \sqrt{\frac{N^2(\sigma_c^2 - \sigma^2)}{Nn_1 - n_1^2}} \quad (4)$$

Finally recall that $p_1 = n_1/N$, so that substituting with $n_1 = p_1N$ gives:

$$|\theta_1 - \theta_2| = \sqrt{\frac{(\sigma_c^2 - \sigma^2)}{p_1(1 - p_1)}} \quad (5)$$

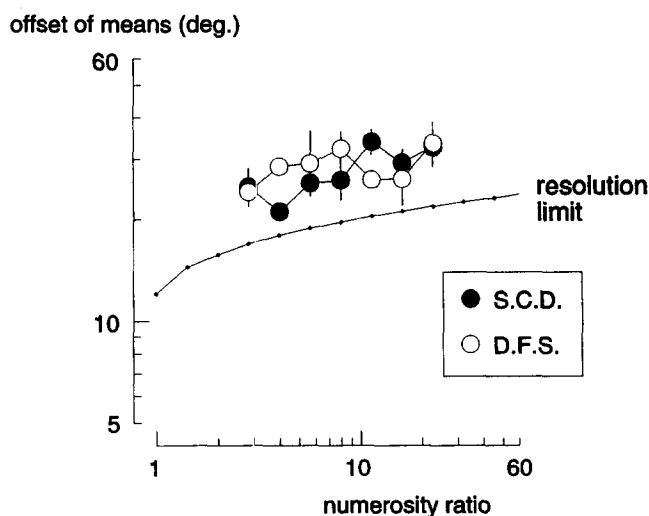
which does not depend on N .

Substituting $\sigma_c^2 - \sigma^2$ in Eq. (5) with the discrimination threshold standard deviation of a single set gives a prediction of the mean offset of two sets which produces this cue in standard deviation. These predictions are plotted for a variety of relative set numerosities in Fig. 4. Note that no further fitting has been used. Performance in this condition is consistent with subjects using variability of orientation to perform the task.

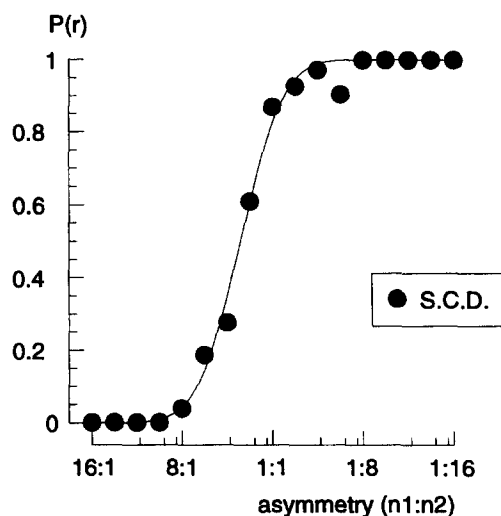
The analysis above assumes that it is the relative number of elements in each set which determines subjects' performance on the task. To determine whether this is correct, a control condition was run. In the stimuli used, each set contained equal numbers but the total number varied from 16 to 512 texels. Data in Fig. 5 show that within the range tested, total number has no appreciable effect on performance. For all subsequent experiments a total of 64 texture elements was employed and pilot studies indicated that there was little difference in performance on any of the tasks when using 256 elements.

Experiment 2: Judging the relative orientations of mixed texel sets

Subjects could perform the task in Experiment 1 based purely on the degree of orientational variability. How can one measure the minimum separation of sets in the orientation domain at which they are resolved? One method is to use a similar task to that used in Experiment 1 but to match the perceived orientation variability. This may not be straightforward if perceived variability is not directly related to a simple orientation statistic such as orientation standard deviation. In this experiment we use a different approach by getting subjects to make a judgement about the relative position, in the orientation domain, of the two sets.



(a)



(b)

FIGURE 6. Threshold offsets for reporting the orientation of the less numerous set relative to the more numerous set, as a function of relative number. Note that all thresholds exceed the theoretical resolution limit. (b) Psychometric function for the same judgement with a fixed offset of set means ($\delta\theta = 45$ deg).

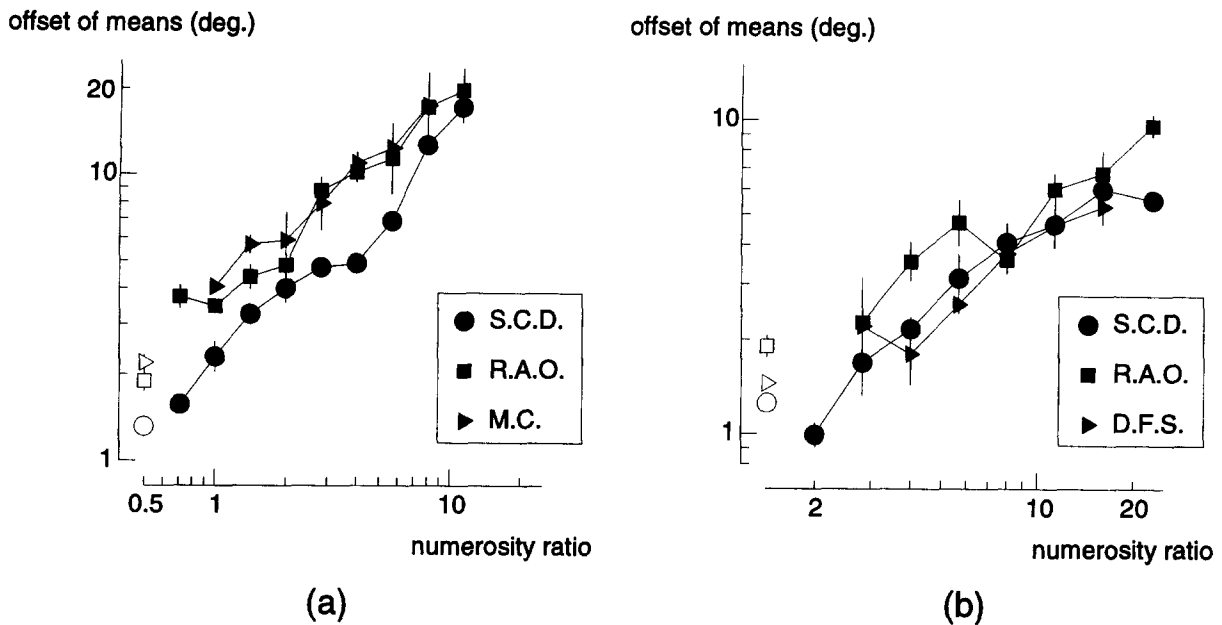


FIGURE 7. Threshold offset of means for three subjects from (a) Experiment 3; and (b) Experiment 4. The unfilled data points are the threshold mean offsets for judgement of average orientation of a single-set texture.

A texture patch composed of two sets containing unequal numbers of texels was presented. Subjects reported whether the less numerous set of texels was clockwise or anti-clockwise to the other set. This is a numerosity asymmetry task. This task used a single interval and to prevent the subjects exploiting the mean of the distributions as a cue, the mean orientation of the more numerous set was randomized over a 180 deg range. Other experimental details were identical to those given in the Methods section, with the exception that a longer exposure duration of 500 msec was used, since pilot trials indicated that subjects were unable to obtain reliable thresholds at shorter exposure times. Subjects were given considerable practice on this task before data collection commenced.

The results [shown in Fig. 6(a)] showed that subjects appear to require resolution of distributions to perform the task. This suggests that relative numerosity information about distributions is only derived if those distributions are distinct perceptual entities separated by at least the resolution limit. The only way subjects could have performed this task without resolving the components would be if they could determine the degree of skewness of the texture orientation distribution. The fact that they cannot suggests that although subjects can accurately assess the first and second moments of orientation distributions they cannot assess the third moment.

Comparison with the last experiment is difficult. In Experiment 2 subjects determined which set was which (according to the number of texels in each) and judged the mean orientation of one in relation to the other. The fact that performance in this task approaches the resolution limit suggests that elevation of thresholds, compared with Experiment 1, is due to the offset of the means of the two orientation distributions. However, it is

possible that the increased complexity of the numerosity judgement was responsible for poorer performance. To confirm that it is the offset of mean orientations which determine performance, a control experiment was performed. The task was a numerosity judgement at a fixed offset of mean orientations ($\delta\theta = 45$ deg, the maximum cue size) for different ratios of numerosity. This experiment will show the subjects' performance as a function of the difference in number between the two sets, but in the absence of any interference due to orientation.

Results as a function of relative numerosity of sets are shown in Fig. 6(b), where a psychometric function for the first author is given. The abscissa is the ratio of the number of elements in the two sets. Asymmetries from 16:1 to 1.4:1 indicate that the mean orientation of the set with more elements was clockwise of the other set. Asymmetries from 1:1.4 to 1:16 indicate that the mean orientation of the set with more elements was anti-clockwise of the other set. Data from this control were collected using a method of constant stimuli (1024 presentations) and fitted with a cumulative Gaussian. It is clear that any effects of absolute difficulty of the numerosity judgement (in isolation from orientation interference) should be insignificant for ratios greater than 1:4 [the standard deviation of the psychometric function shown in Fig. 6(b) is 1:2.1].

EXPERIMENTS 3 AND 4: JUDGING THE OVERALL ORIENTATIONS OF MIXED TEXEL SETS

The next two experiments investigated the extraction of average orientation from texture and tested the predictions of a number of localization models, adapted from the spatio-luminance to the orientation domain.

Subjects were presented with one texture field and asked to indicate whether the average orientation was clockwise or anti-clockwise relative to vertical. In Experiment 3 the field was composed of two sets with mean orientations of 90 deg and 90 deg + $\delta\theta$, respectively. The offset of the means of the two sets ($\delta\theta$) was systematically varied. Additionally, thresholds were determined for comparing the mean orientation of a single set with vertical. All textures were presented for 100 msec.

Figure 7(a) shows the data from this experiment. Thresholds increased with the ratio of the number of elements in the two sets and were consistently lower than thresholds from the resolution and asymmetry tasks. The thresholds for the single set stimuli are shown as unfilled data points. Notice that thresholds for the single set stimuli are all between 1.2 and 2.5 deg. These values are not much higher than orientation thresholds reported for line and grating stimuli (0.47–1.5 deg, Heeley & Buchanan-Smith, 1990; Heeley & Timney, 1988; Matin & Drivas, 1979; Matin *et al.*, 1987; Vogels & Orban, 1985). This is surprising given that they are composed of spatially disorganised patches with an orientation standard deviation of 6.0 deg, but is consistent with earlier work on orientation discrimination of DoG patches in the presence of orientation perturbation (threshold of 0.78 deg for a perturbation of 3 deg; Paradiso *et al.*, 1989).

In order to better separate the predictions of the models described in the next section a larger asymmetry was set up in Experiment 4, using three distributions. This experiment should also verify that whatever statistic was being used to perform the task, generalizes to textures composed of more than two resolvable distributions. The probability of a particular orientation was:

$$p(\theta) = \frac{1}{\sigma\sqrt{2\pi}} \left[p_1 \exp \left[-\frac{(\theta - \theta)^2}{2\sigma^2} \right] + p_2 \exp \left[-\frac{(\theta - \theta - \delta\theta)^2}{2\sigma^2} \right] + p_2 \exp \left[-\frac{(\theta - \theta - 3\delta\theta)^2}{2\sigma^2} \right] \right]$$

where $p_2 = (1 - p_1)/2$ and all other variables are identical to those given for Eq. (1). Again the threshold $\delta\theta$ is for the discrimination of average orientation of a texture from vertical.

Threshold mean orientation offsets are shown in Fig. 7(b). Note that the numerosity ratios shown are based on the ratio of the number of elements in one of the flanking distributions to the number in the more numerous set (p_1/p_2). Unfilled symbols show the thresholds for the orientation judgement of a single set, as described above. Thresholds show the same monotonic dependence on numerosity ratio and are again lower than thresholds from the asymmetry and resolution tasks.

COMPARISON OF FEATURES FOR CODING MEAN ORIENTATION

In this section four schemes for coding the central

tendency of an orientation distribution are examined. Consider a compound distribution of texel orientations with a particular ratio of the number of elements in each set, and an offset of component sets equal to the threshold value for that stimulus (as determined from Experiments 3 and 4). Features of this distribution, e.g. the mean and the peak, will not coincide. The difference between the orientation of the candidate feature and vertical is assumed to be the cue that the subject uses. The threshold shift to judge the average orientation of a single set provides a comparison. If subjects use a particular cue to judge average orientation of the compound textures, then the threshold shift of this cue should equal the average orientation threshold for the single set judgement.

Four candidate features were tested: peaks, threshold edges, zero-crossings and centroids. All features were extracted from idealized orientation distributions.

- Zero-crossing: the zero crossing model assumes convolution of the idealized orientation distribution with a smoothed second differential operator. This is similar to the “orientational filters”, proposed by Keeble *et al.* (1995) as a mechanism for discriminating oriented textures. It is also similar to a component of Paradiso’s (1988) model of orientation perception, where the output of cells within an orientation column are filtered with a Difference-of-Gaussian filter.

In this case a Laplacian-of-Gaussian operator was used, of the form:

$$\nabla^2 G(x, \sigma) = \frac{1}{\sigma^2} \left(\frac{x^2}{\sigma^2} - 1 \right) e^{-x^2/2\sigma^2}$$

Because orientation is cyclical, convolution was wrapped around 180 deg. Location was then defined as the mid-point between the outermost two zero-crossings of the convolution about vertical. The width of filter employed was the value minimizing the Chi-square of the fit of the model to the data from Experiment 3. This value was a filter standard deviation of around 6 deg (mean = 6.10 deg, SD = 0.81 deg).

- Threshold edge: two points in the distribution were located which were equal to some threshold value. Location was defined as the mid-point between these points. The threshold employed minimized the Chi-square of the fit of the model to the data.
- Peak: the location of a distribution is assigned to the maximum of the idealized orientation distribution.
- Centroid: the centroid was calculated using:

$$\theta = \begin{cases} \theta_c & \text{if } -2 \left[\sum_{i=1}^{180} p(\theta_i) \cos 2(\theta_i - \theta_e) \right] \leq 0 \\ \theta_e + 90^\circ & \text{otherwise} \end{cases}$$

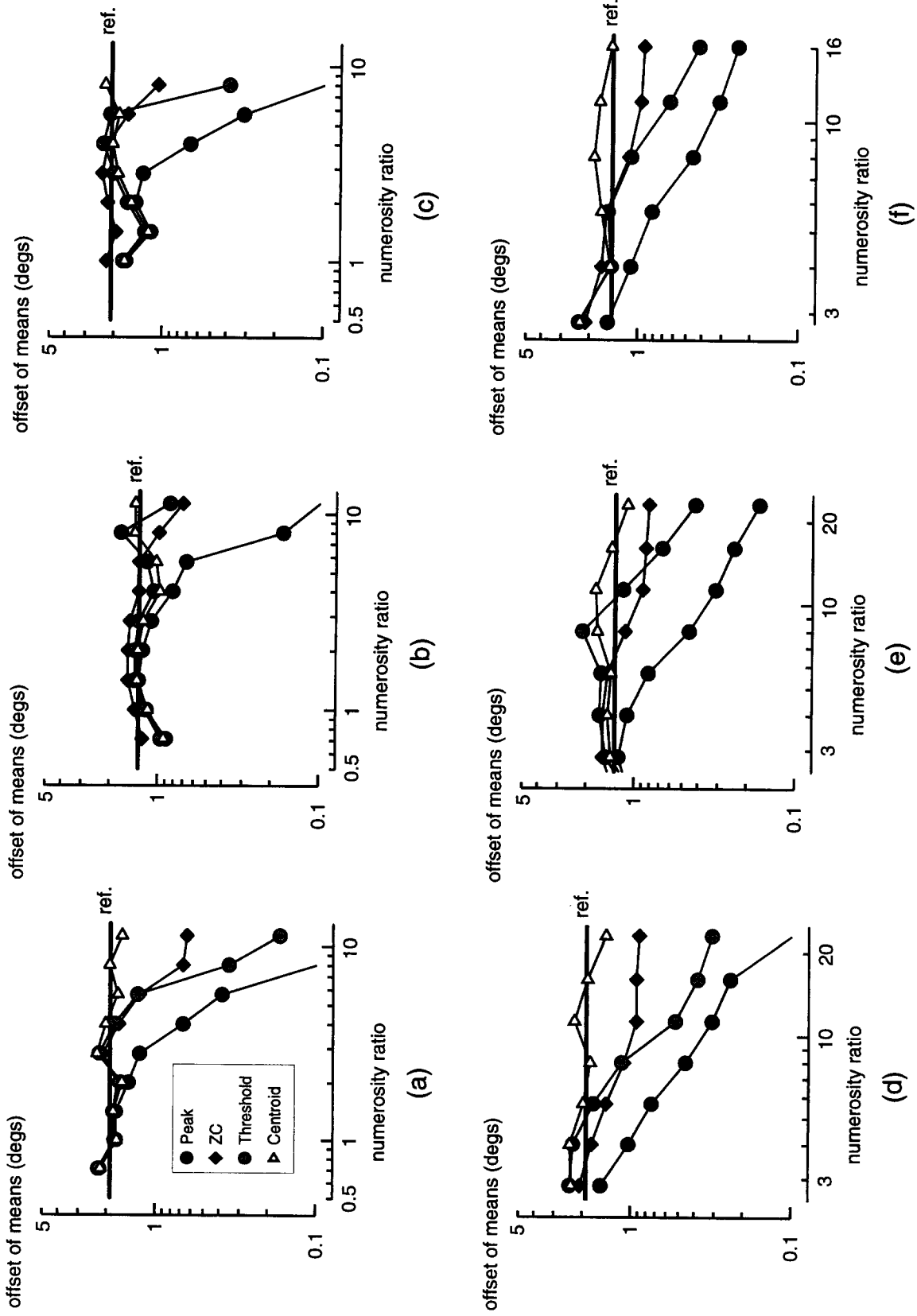


FIGURE 8. Subjects' performance on the global orientation judgement re-expressed as shifts in different features in the idealized orientation distribution. (a–c) Data from Experiment 3 (Subjects RAO, SCD and GP, respectively). (d–f) Data from Experiment 4 (RAO, SCD and MC, respectively).

where:

$$\theta_e = \frac{1}{2} \tan^{-1} \left[\frac{\sum_{i=1}^{180} p(\theta_i) \sin 2\theta_i}{\sum_{i=1}^{180} p(\theta_i) \cos 2\theta_i} \right]$$

(The derivations of these formulae are given in Dakin, 1997). Note that integration occurs over the full range of orientations. In practice integration over a restricted range of orientations can be more appropriate (e.g. when a distribution contains two resolved entities).

Given a threshold difference in the mean orientations of the component sets, the size of the cue given by each of the candidate features may be calculated. These values should be equal to the size of the cue for the single-set texture at threshold offset. Cue sizes based on thresholds from Experiment 3 are shown in Fig. 8(a–c), and those based on thresholds from Experiment 4 are shown in Fig. 8(d–f).

Typically the models produce cues which are smaller than the threshold shift in central tendency, even given the idealized input, and can therefore be eliminated. The peak model in particular fails to explain the data. The best fit is given by the centroid model, although the threshold edge and zero-crossing model are also adequate. Note that the latter two models are fitted, whereas the centroid model has no free parameters.

For the simulation of Experiment 4, the same threshold values and filters were used (except in the case of subject DFS where no data from Experiment 3 were available, and these values were derived using the data from this experiment). Figure 8(d–f) demonstrates that the increase in asymmetry of the stimuli in Experiment 4 further separates performance of the four models. Again the centroid model best explains the data.

These results suggest that the centroid of a set of orientation measures is the best model of human judgement of the central tendency of a set of orientation estimates made from a texture.

DISCUSSION

The main findings from the experiments described in this paper are as follows:

- Subjects can discriminate stimuli composed of two oriented texel sets from a field composed of one set at mean orientation offsets which depend on the relative number of elements in each set (Experiment 1).
- The global orientation variation of a stimulus can be used to discriminate textures composed of multiple sets from those containing a single set. As a consequence, threshold orientation offsets of multiple sets can be below the theoretical resolution limit (Experiment 1).

- When the orientation variation cue is eliminated, by using a relative numerosity judgement, thresholds are elevated above the theoretical resolution limit. Information about the skew of orientation distribution is therefore not available. The resolution limit gives a lower limit on when texel sets become separate, in the sense that properties of those sets can be measured independently (Experiment 2).
- Subjects accuracy at estimating the orientations of textures composed of skewed distributions of orientations is best accounted for by a model extracting the centroid of orientations (Experiments 3 and 4).

The centroid model as presented makes two assumptions: firstly, that orientation measures are extracted from the whole image and secondly, that integration over the whole range of orientations should be used to make the estimate. Both assumptions are certainly false, since we are quite capable of perceiving orientation structure in natural images, requiring multiple local orientation estimates, and we can perceive structure in multi-modal textures. In Experiment 2, subjects clearly have access to relative numerosity information about two texel sets and an orientation primitive based on some portion of the orientation distribution is suggested. Although the zero-crossing model produces a reasonable fit to the data from Experiments 3 and 4 it was also run (with identical filter characteristics) on the data from Experiment 1 to yield predictions of orientation variation discrimination thresholds. The model produced adequate predictions at small ratios but did not explain the data when the number of elements in one set became small. Small distributions are smoothed out entirely. This challenges the assumption that an idealized orientation distribution is available.

In reality, the visual system estimates the orientation statistics of an observed texture by integrating a set of local estimates of orientation. This could occur through spatial summation of similarly oriented receptive field outputs within a “region of integration”. Using such a scheme the relative output of differently oriented mechanisms at a point could signal the reliability of extracted features. What are the spatial dimensions of this region of integration? In order to achieve scale invariance one would expect its size to scale with the spatial frequency of filters used to derive local orientation. Kingdom *et al.* (1995) have found this to be the case for textures containing orientations modulated over space.

Textures with spatially modulated orientation structure have been used previously to explore the analogy between the processing of orientation and luminance distributions. Keeble *et al.* (1995) had subjects discriminate between sinusoidally modulated, and uniform orientation p.d.f.s, as a function of the amplitude and frequency of the sinusoid. Subjects perform optimally with distributions modulated at one cycle per 180 deg. Keeble *et al.* (1995) modelled the task using an “orientational filter” with an average half-height, full-width of 34 deg. This is very large compared with the

filter sizes derived from the average orientation estimation task, but is probably attributable to task differences. In Keeble *et al.* (1995) subjects had only to report which texture was the more anisotropic. In Experiments 3 and 4 the threshold mean orientation offsets of 1.0–2.0 deg suggests a much finer-scale filter is appropriate. In the case of the Keeble *et al.* (1995) task, such a filter would probably respond to spurious orientational structure, rather than the orientational modulation subjects were required to detect.

Examining a typical stimulus for higher ratios as shown in Fig. 3(d, e), it is apparent that, as the numerosity ratio of the two distributions increases, the judgement from Experiment 1 becomes closer to a “search” task, as described in the literature on visual attention (e.g. Treisman & Gormican, 1988). In Experiment 1 it appears that, in the case of multiple targets, threshold orientation difference between target and distractor may be well modelled by predictions based on change in the orientation standard deviation of the texture. It is interesting to note that, for example, such a model would predict a non-monotonic dependence of performance upon distractor number (shown by, for example, Sagi, 1990; Sagi & Julesz, 1987). From statistical considerations a small total number of texels leads to a poor estimate of orientation standard deviation, producing poor performance which improves with increasing texel number. However, as number increases, the influence of a single distractor on the global standard deviation estimate will decrease. Subjects performance may be determined by the total of these two sources of error. The suggestion is that the underlying statistics of texel orientation may explain some visual search effects without recourse to specific, spatial interactions.

In conclusion, the data reported strongly constrain a representation of texture orientation statistics. The variability and centroid of an orientation distribution are available to subjects, but skew information is not. Multiple sets can be perceived and independent information about them is represented when they are resolved. These results suggest a parallel between the visual processing of spatially unstructured texture and spatio-luminance information.

REFERENCES

Badcock, D. & Westheimer, G. (1985). Spatial localization and hyperacuity: the centre/surround localization contribution function has two substrates. *Vision Research*, *25*, 1259–1267.

Barlow, H. (1980). The absolute efficiency of perceptual decisions. *Philosophical Transactions of the Royal Society of London B*, *290*, 71–82.

Blake, A. & Marinos, C. (1990). Shape from texture: estimation, isotropy and moments. *Artificial Intelligence*, *45*, 323–380.

Dakin, S. C. (1997). The detection of structure in Glass patterns: psychophysics and computational models. *Vision Research*, *37*, 2227–2246.

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

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

Heeley, D. W. & Timney, B. (1988). Meridional anisotropies or orientation discrimination for sine wave gratings. *Vision Research*, *28*, 337–344.

Hubel, D. & Wiesel, T. (1967). Receptive fields, binocular interaction and functional architecture in the cats' visual cortex. *Journal of Physiology*, *160*, 106–154.

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

Kass, M. & Witkin, A. (1985). Analyzing oriented patterns. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, (pp. 944–952). Los Angeles, CA.

Keeble, D. R. T., Kingdom, F. A. A., Moulden, B. & Morgan, M. J. (1995). Detection of orientationally multimodal textures. *Vision Research*, *14*, 1991–2005.

Kingdom, F. A. A., Keeble, D. & Moulden, B. (1995). Sensitivity to orientation modulation in micropattern-based textures. *Vision Research*, *35*, 79–91.

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

Marr, D. (1982). *Vision*. San Francisco, CA: Freeman.

Matin, E. & Drivas, A. (1979). Acuity for orientation measured with a sequential recognition task and signal detection methods. *Perception and Psychophysics*, *25*, 161–168.

Matin, E., Rubsamen, C. & Vannata, D. (1987). Orientation discrimination as a function of orientation and spatial frequency. *Perception and Psychophysics*, *41*, 307–404.

Nothdurft, H. C. (1991). Texture segmentation and pop-out from orientation contrast. *Vision Research*, *31*, 1073–1078.

Paradiso, M. (1988). Theory for the use of visual orientation information which exploits the columnar structure of striate cortex. *Biological Cybernetics*, *58*, 35–49.

Paradiso, M. A., Carney, T. & Freeman, R. D. (1989). Cortical processing of hyperacuity tasks. *Vision Research*, *29*, 247–254.

Phillips, G. & Wilson, H. (1983). Orientation bandwidths of spatial mechanisms measured by masking. *Journal of the Optical Society of America*, *62*, 226–232.

Sagi, D. (1990). Detection of an orientation singularity in Gabor textures: effect of signal density and spatial-frequency. *Vision Research*, *30*, 1377–1388.

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

Treisman, A. & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries. *Psychological Review*, *95*, 15–48.

Vogels, R. & Orban, G. A. (1985). The effect of practice on the oblique effect in line orientation judgements. *Vision Research*, *25*, 1679–1687.

Vorhees, H. & Poggio, T. (1987). Detecting textons and texture boundaries in natural images. In *Proceedings of the First International Conference on Computer Vision* (pp. 250–258).

Watt, R. J. (1991). *Understanding vision*. Academic Press, London.

Watt, R. J. & Andrews, D. (1981). APE: Adaptive probit estimation of psychometric functions. *Current Psychological Review*, *1*, 205–214.

Watt, R. J. & Morgan, M. J. (1983). Mechanisms responsible for the assessment of visual location: theory and evidence. *Vision Research*, *23*, 97–109.

Whitaker, D. & Walker, H. (1988). Centroid evaluation in the vernier alignment of random dot clusters. *Vision Research*, *28*, 777–784.

Wilson, H. & Gelb, D. (1984). Modified line-element theory for spatial-frequency and width discrimination. *Journal of the Optical Society of America A*, *1*, 124–131.

Witkin, A. (1981). Recovering surface shape and orientation from texture. *Artificial Intelligence*, *17*, 17–47.

Wolfson, S. S. & Landy, M. S. (1995). Discrimination of orientation-defined texture edges. *Vision Research*, *35*, 2863–2877.

Acknowledgements—The authors would like to thank Ian Paterson for many comments and suggestions. This work was funded by the SERC (grant GR/H53181).