Pergamon

PII: S0042-6989(97)00038-2

Vision Res., Vol. 37, No. 16, pp. 2227–2246, 1997

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

The Detection of Structure in Glass Patterns: Psychophysics and Computational Models

STEVEN C. DAKIN*

Received 2 October 1996; in revised form 19 December 1996

Experiments are reported which examine the judgement of the mean orientation of textures composed either of short lines or dipoles (*Glass patterns*). The effects of element length, density, and orientation variation are described. Psychophysical data are compared with predictions from four schemes for extracting features from Glass patterns: token matching, isotropic filtering, oriented filtering, and "adaptive" filtering (selection of local peak output from multiply oriented filters). Glass patterns are spatially broadband but only contain orientation structure at a narrow range of scales making them suitable for examining how *filter size* is selected for texture processing. A criterion for scale selection is proposed: that local variation of feature orientation should be minimized. Simulations indicate that neither models using isotropic filtering nor token matching achieve human levels of performance on certain tasks. Adaptive filtering, operating at a scale selected using the criterion described, provides good agreement with the psychophysical data reported and is a practical scheme for deriving features using oriented filters. © 1997 Elsevier Science Ltd.

Grouping Orientation Scale Texture

INTRODUCTION

The present study examines two issues associated with how visual features are extracted using spatial filters. The first is whether oriented or isotropic filters are used. The second is how information is selected from, or combined across, filters of different sizes and (if filters are oriented) orientations. In this paper these problems are considered in the context of the perception of structure in a class of texture known as Glass patterns.

Glass patterns

Glass patterns are composed of the superimposition of one or more copies of a field of randomly distributed features (e.g., dots) onto the original, where the copy is a geometric transformation of the original [Fig. 1(a); Glass, 1969; Glass & Perez, 1973). The visual impression is of a compelling oriented structure with dot pairs (*dipoles*) aligned along the direction of the local transformation. Glass patterns are interesting for a number of reasons. Firstly, they approximately isolate the selection of orientation from other visual processes owing to, e.g. contrast and size differences (Zucker, 1982). Secondly, they contain structure at only a limited range of spatial scales, which makes them ideal for investigating scale selection/combination processes (see below). Finally, although a number of computational and psychophysical investigations of the perception of Glass patterns have been conducted, none of the models proposed has been shown to be completely satisfactory.

The problem of deriving structure from Glass patterns is closely linked to a key problem for computational accounts of texture perception, and that is how features from natural texture are extracted and represented. Within the psychophysics literature, this problem is often side-stepped by using textures composed of spatially distinct micro-patterns, which are clearly delineated by their brightness. However, considering Fig. 1(b), which shows a highly oriented natural texture, no such features present themselves. Spatial filtering has been proposed as a computationally efficient method both for deriving features from visual texture (Vilnrotter *et al.*, 1986; Vorhees & Poggio, 1987, 1988; Wen & Fryer, 1991) and as a mechanism for grouping in Glass patterns (e.g. Kass & Witkin, 1985; Zucker, 1982). Figure 1(c–h) shows the thresholded response of Laplacian-of-Gaussian (LoG) filters to the Glass pattern and the tree bark texture. It is clear that the dark blobs in the output of the medium scale filter [Fig. 1(e, f)] reflect the orientation of dipoles, and the local orientation of the bark texture. This is not the case for all scales and Prazdny (1986) has pointed out that a particular problem for any account of grouping in Glass patterns using filtering is just how filter size is selected. In the past, models have either made assumptions about the setting of filter size (Kass & Witkin, 1985; Zucker, 1982), or neighbourhood size for token matching (Stevens,

*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].

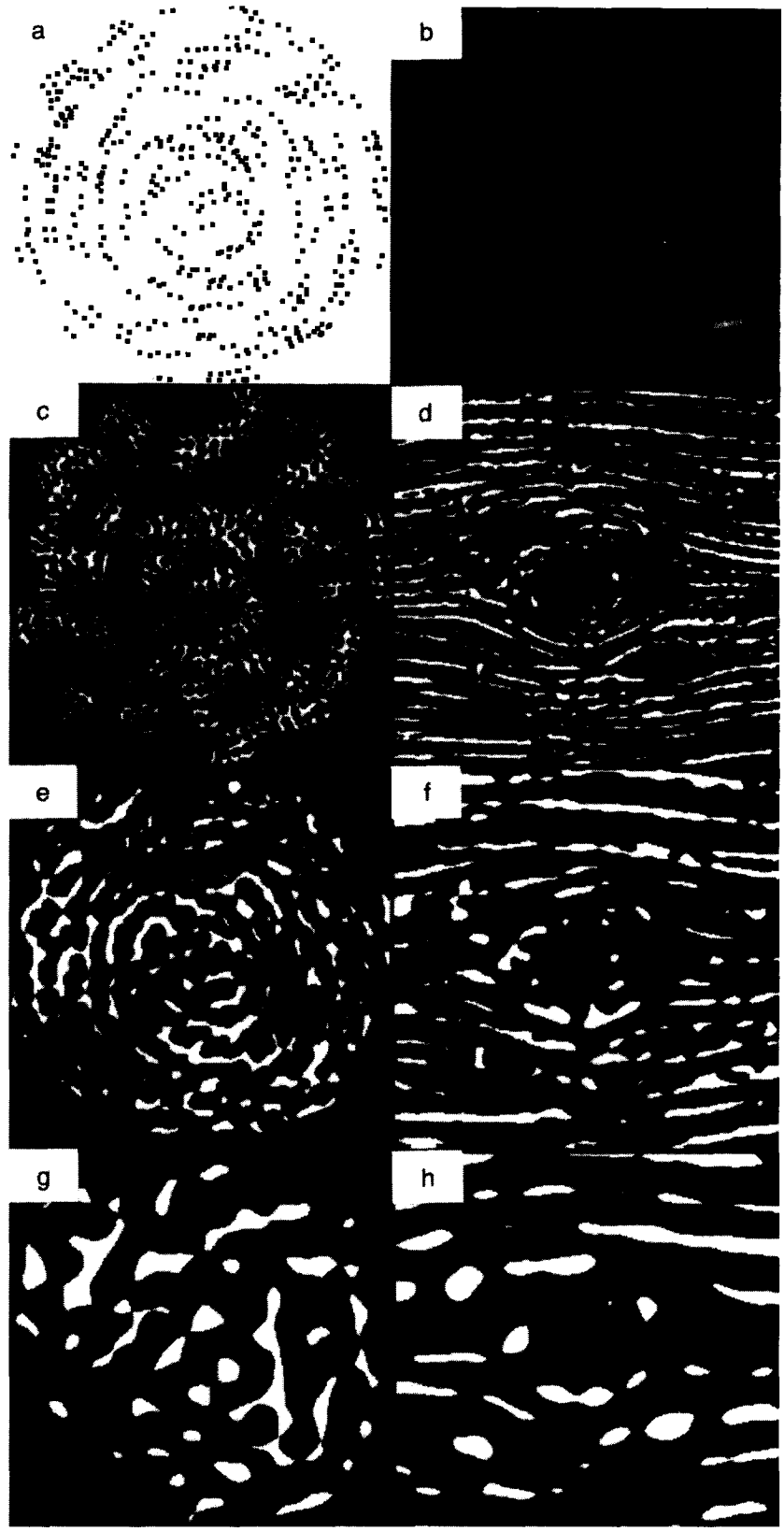


FIGURE 1. Locating features in Glass patterns and natural textures: (a) rotational Glass pattern; (b) tree bark (texture d72 from Brodatz, 1966). (c–h) Response of isotropic, Laplacian-of-Gaussian filters to (a) and (b), with space constants of (c, d) 2, (e, f) 4, and (g, h) 8 pixels. (Grey levels above or below a threshold value have been replaced with white or black pixels, respectively, to highlight "blobs".)

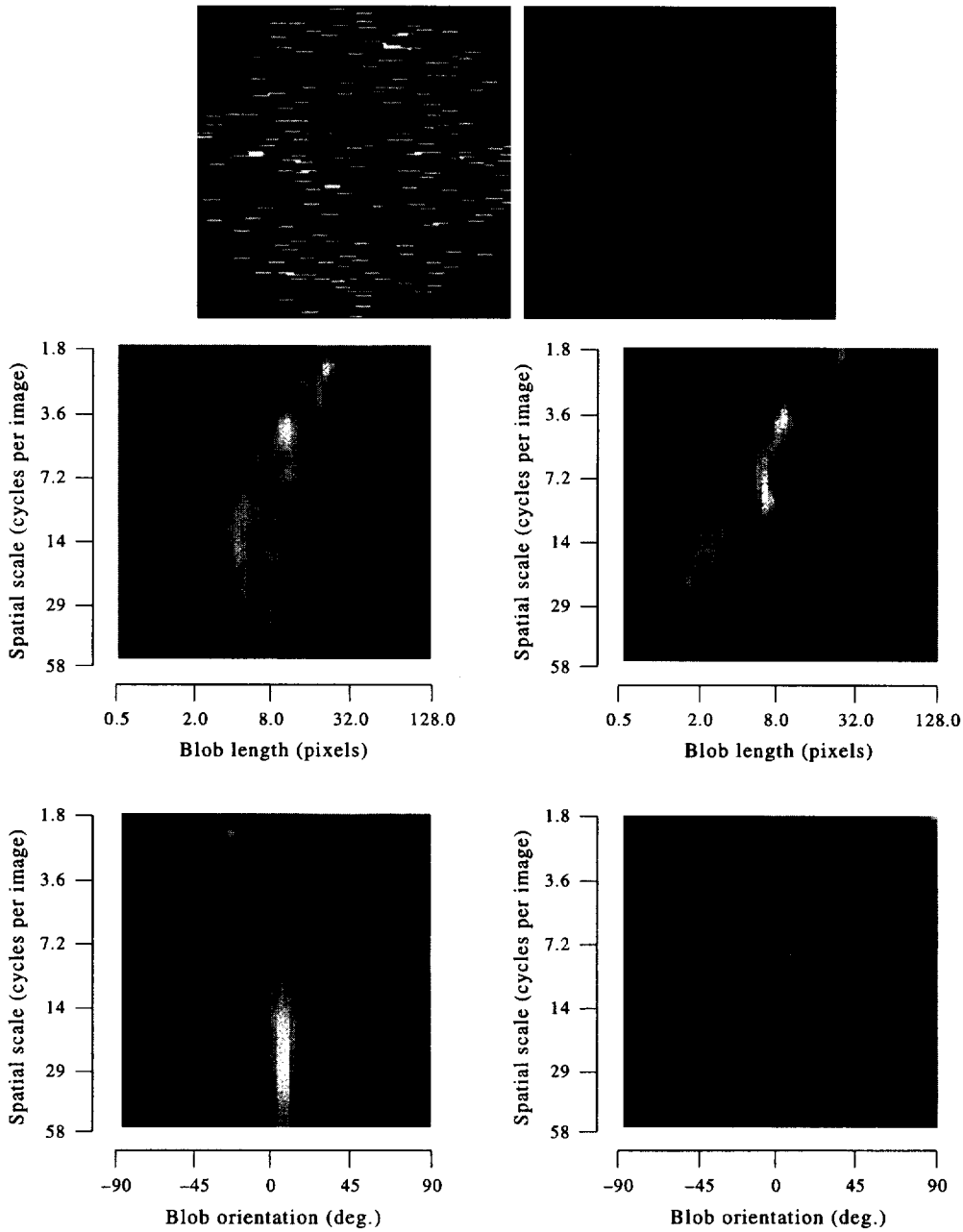


FIGURE 2. Top row: A line texture and a Glass pattern with similar mean orientation, element length and number. Scale–length histograms (middle row) and scale–orientation histograms (bottom row) of Laplacian-of-Gaussian features derived from the line texture and Glass pattern. Brightness indicates the number of blobs in the output of a filter at a particular orientation or length.

1978) or cannot explain global organization effects at all (e.g., the autocorrelation model of Maloney *et al.*, 1987). Prazdny (1986) concludes that there must be "an evaluating agent" looking at the output of the filters at various scales which, he suggests, is not unlike the Gestalt notion of "Prägnanz".

Figure 2 further illustrates the importance of filter-size selection for Glass patterns. It shows histograms* of a Glass pattern and a line texture illustrating how orientation and length statistics change with the spatial scale of analysis. The narrow, vertical band in the line texture scale–orientation histogram (bottom row, left) indicates that a wide range of fine spatial scales will give an accurate estimate of the correct orientation of the line texture. In contrast, the "hour-glass" shape of the histogram derived from the Glass pattern (bottom row, right) shows that the range of orientations present in the pattern is determined by the filter size selected. Notice that the length–scale histogram of the Glass pattern indicates that mean blob length increases with scale. This information does not appear to be useful for the selection

*Histograms were derived using symbolic blob descriptions (Watt, 1991), derived from the half-wave rectified outputs of a range of sizes of LoG filters. For further details of this process see the "Modeling" section below.

of scale. Indeed the primacy of orientation information in the processing of these patterns has been demonstrated psychophysically by Caelli & Julesz (1979), who showed that variance of orientation, but not length, determines the discrimination strength of patches of dipoles.

Of course, defining the "correct" scale is not possible in the absence of a particular visual task. Let us assume that the goal of the process responsible for deriving structure from oriented texture is to maximize the accuracy of its estimates of local orientation. Statistical wisdom indicates that the source of information it uses should have minimal variance. This suggests a strategy for determining any free parameters (such as filter size) of a texture processing model; set them so that derived features have minimal orientation variance. By examining a symbolic image description, the local orientation variance of a given set of blobs may be calculated, and so an estimate of the reliability of a particular spatial scale for estimating local orientation may be assessed (see the Appendix for details of the variance calculation). Referring back to Fig. 2, this is equivalent to selecting the scale at which the scale–orientation histogram is narrowest.

The psychophysical task used in this paper is the judgement of the mean orientation of oriented texture patterns. The more common "structure vs no-structure" task was not used because it does not sufficiently constrain the source of information used by the subject (they could use any form of regularity in the patterns). The requirement of an accurate estimate of mean orientation constrains the subjects' behaviour in a way that may be built into a model directly, in the manner described in the preceding paragraph. The remainder of this paper presents three experiments examining the extraction of the mean orientation of Glass patterns as a function of a number of stimulus parameters. In order to isolate how the grouping of dots into dipoles affects the task, performance with textures composed of lines is also measured. Differences between line and dipole textures should be attributable to grouping uncertainty. Finally, four models are used to simulate performance on these tasks. Three used spatial filtering in conjunction with thresholding and symbolic feature description. The filtering used was either Laplacian-of-Gaussian, oriented Difference-of-Gaussian (DoG), or "adaptive" which combined the outputs of DoGs at multiple orientations. In each case, filter size was selected by minimizing the variability of the orientation of derived features. The final model was a token matching scheme specifically designed to derive dipoles from Glass patterns by maximizing the parallelism of local dot matches within a local neighbourhood (Stevens, 1978).

GENERAL METHODS

The following three experiments manipulated different parameters of the line and dipole textures but used the same basic method. The independent variables examined were the length of elements (Experiment 1), the standard deviation (SD) of the orientation of elements (Experiment 2) and the density/number of elements (Experiment 3). Experiments 1 and 2 compare data from tasks using line and Glass pattern textures, while Experiment 3 examines only Glass patterns. The subjects' task in all experiments was to judge the mean orientation of the texture presented.

Subjects

Three subjects served as subjects in the experiments. All were experienced in psychophysical procedure, and FJM and RAO were naïve to the purpose of the experiments. All subjects had normal or corrected-to-normal vision and undertook sufficient practice to reach asymptotic performance before threshold measurement began.

Apparatus

The generation and presentation of stimuli, and the recording of subject responses was carried out on a Macintosh IIfx microcomputer. The display was a Formac ProNitron 80.21 colour monitor with a frame refresh rate of 75 Hz. The screen was viewed binocularly with natural pupils at a distance of 2 m.

Stimuli

All stimuli used in the experiments were approximately circular texture fields, with radii subtending 1.23 deg (128 pixels), of either lines or dot pairs (dipoles). These fields appeared within a 2.46 deg (256 pixel) square image. Lines or dipoles appeared white on a black background, and were distributed randomly throughout the field.

Line elements were anti-aliased, using 16 grey levels. Component dots of the dipoles were individual pixels which subtended approximately 35 arc sec. No anti-aliasing of dipoles was used. If the orientation of a dipole required that one of the component dots be placed in a position between the discrete pixel locations available, the nearest pixel location was used. As a consequence, in the experiments employing a constant dipole length (Experiments 2–3), a relatively large value of 8 arc min (approximately 14 pixels) was used (see Fig. 4 for examples). At this length, cues as small as 4 deg may be presented.

The orientation of elements of both line and dipole textures were drawn from Gaussian random distributions (clipped at $\pm 6\sigma$). Apart from Experiment 2, where the effect of orientation variation was investigated, distributions had a SD of 8 deg. This value was used because pilot studies indicated that such a level of variability, with textures composed of 8 arc min long dipoles, reduced performance (by around 50% from conditions with no variation in element orientation), bringing thresholds closer to the minimum cue which could be reliably presented. Vertical patterns were chosen to avoid problems due to the well known oblique effect (e.g. Appelle, 1970), and because of the established advantage for vertical over horizontal Glass patterns (Jenkins, 1985).

Textures in Experiments 1 and 3 were composed of 512 lines or dipoles, corresponding to an average density of 89.4 elements/degree^2 (Experiment 3 explicitly investigated the effect of element number/density). A pilot study indicated only a small effect of element density on thresholds for the line and dipole textures. Relatively dense patterns were used because theories based on symbolic matching of tokens (Stevens, 1978) predict that performance should be poor under these conditions.

Procedure

Stimuli were presented in the centre of the screen, which was indicated by a pre-stimulus marker, for 100 msec. An ISI of 750 msec followed each response.

The subjects' task was a single interval, two-alternative forced-choice, and was to report whether the texture presented had an orientation clockwise or anticlockwise of vertical. No reference orientation was presented to subjects. Subjects indicated their response by depressing one of two keys on the computer keyboard.

An adaptive method of constant stimuli, APE (Watt & Andrews, 1981), was used to sample a range of mean orientations around vertical. Three runs of 64 trials were undertaken for each data point presented. Conditions were not interleaved. The psychometric functions measured were the probability of reporting a clockwise orientation as a function of the cue added to the mean orientation (i.e., they measured performance between 0 and 100%). At the end of a block, probit analysis was used to estimate the SD of the psychometric function for each run. The data points plotted are the arithmetic mean of these values, and the error bars show ± 1 SE.

EXPERIMENT 1: EFFECT OF ELEMENT LENGTH ON THE JUDGEMENT OF MEAN ORIENTATION

The first experiment investigated the effect of element length on the judgement of the mean orientation of translational line textures and Glass patterns. Variation of dipole length effectively varies the level of noise in the matching of dipole elements owing to the proximity of other, uncorrelated dots. Subjects' thresholds for line textures give an estimate of the absolute limits on the estimation of the mean of a set of oriented elements in the absence of matching uncertainty.

Previous work examining the effect of dipole length on the perception of Glass patterns has used rating judgements (Caelli & Julesz, 1979) or discrimination of structure from noise (Jenkins, 1983; Wagemans *et al.*, 1993). Caelli & Julesz (1979) used true rotational Glass patterns, i.e., dipole length increases with distance from the centre, and asked subjects to rate how far from the centre they could see structure. They found these estimates of "perceived extent" fall steadily as a function of increasing angle of rotation. Jenkins (1983) determined that discrimination of Glass patterns from noise falls as dipole length is increased and quotes a limit of 1.4 deg for 50% correct discrimination of signal from noise. Similarly, Wagemans *et al.* (1993) measured the discriminability of Glass patterns, composed of 16 dipoles, from noise patterns as a function of dot separation. It was found that increasing dipole separation produces worse discrimination and that length variation was also found to have a detrimental effect on signal detection.

Wagemans *et al.* (1993) and Jenkins (1983) have both interpreted their data as evidence for the models proposed in each. The finding of Jenkins (1983), that there is a critical dipole length for perception of structure, seems to support a matching mechanism using spatial correlation; although virtually any grouping model would also predict such a limit. Wagemans *et al.* (1993) report that the efficiency of subjects at performing the discrimination task (d' as high as 4.6, i.e., 100% correct discrimination) is not matched by the simulated annealing model proposed (maximum of around 85% correct discrimination). Although the model produces "the same rank ordering of performance levels" (Wagemans *et al.*, 1993) this is a weak criterion for assessing the validity of a model. In short, there have been no wholly convincing quantitative explanations of the effect of dipole length on the perception of structure in Glass patterns.

Stimuli

The textures used were composed of 512 lines or dipoles, their orientations drawn from Gaussian random distributions with a SD of 8.0 deg. A range of dipole/line lengths was tested from 1.41 to 32.0 arc min in multiplicative steps of $\sqrt{2}$, and examples of the stimuli are shown in Fig. 3.

Results

Threshold offsets for the mean orientation judgement as a function of element length are shown for three subjects in the lower panel of Fig. 3. The accuracy of judging the mean orientation of line and dipole textures improves rapidly with element length up to 4–5 arc min. This improvement could be due to two factors. Firstly the dependence of the accuracy of orientation estimates on the aspect ratio of the figure (Vassilev *et al.*, 1981; Westheimer, 1981). Secondly, when dipole separation is small the cue has to be relatively large to overcome the problem of the discrete location of pixels comprising each dipole. However this factor does not appear to be a major contributor to poor performance at small element lengths because performance using the line textures, which are anti-aliased and so do not suffer from this problem, closely follow that of the dot textures at short element lengths.

As element length increases above 4–5 arc min, accuracy for judgement of line texture orientation improves until it asymptotes at around 2.0 deg. Judgement of the orientation of Glass patterns, however, quickly breaks down as dipole length increases beyond 8 arc min because of uncertainty in matching the dipole components. The task becomes impossible with dipole textures at around 23–32 arc min for this dot density.

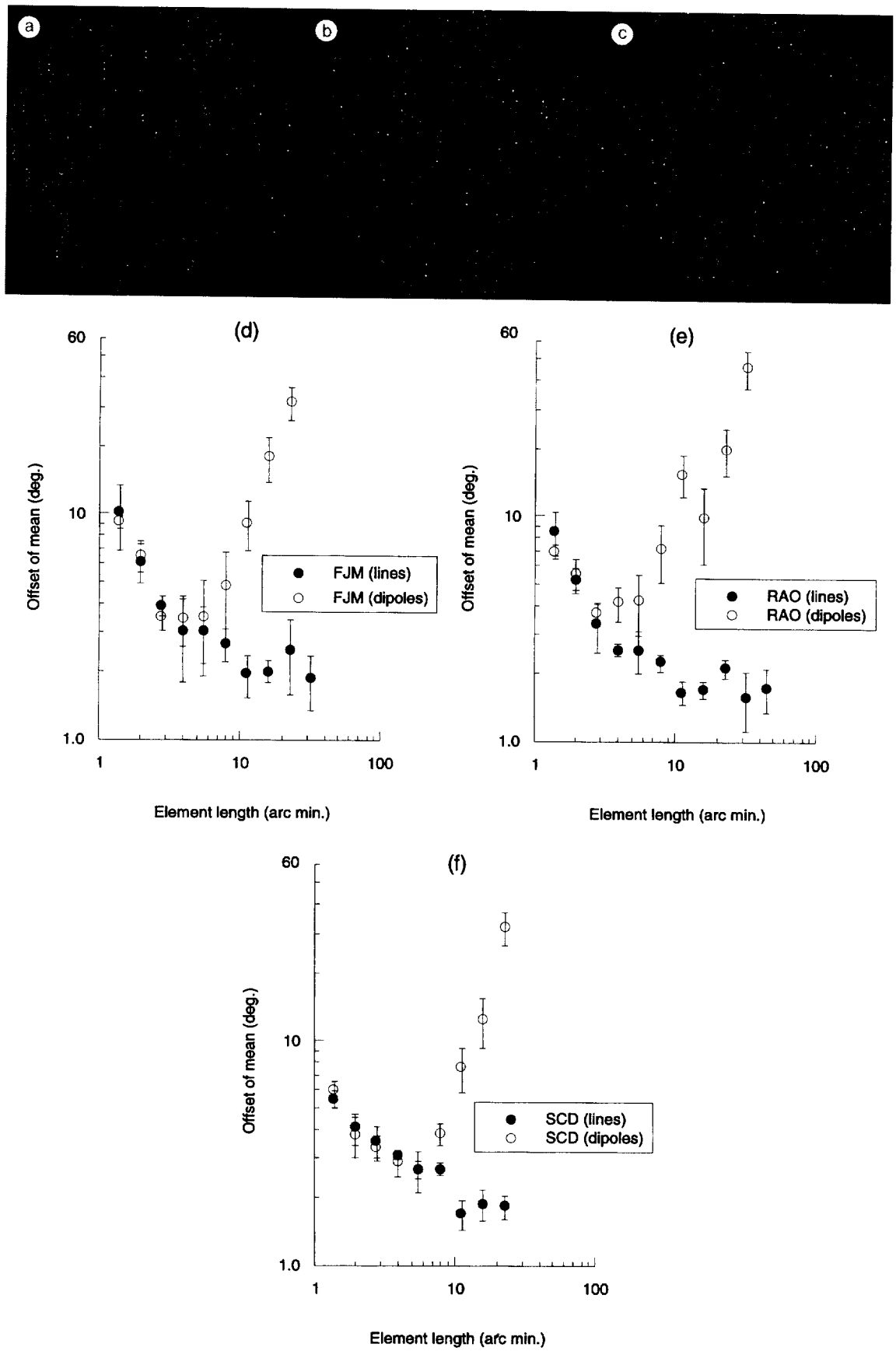


FIGURE 3. (a–c) Examples of the stimuli used in Experiment 1. The patterns shown have dipole lengths of (a) 2.8; (b) 5.6; and (c) 11.3 arc min. (d–f) Threshold mean orientation offsets for judgement of the orientation of line and dipole textures as a function of element length. Results from three subjects are shown.

This deterioration in performance as dipole length increases, confirms the general finding of Wagemans *et al.* (1993) and Jenkins (1983). Experiment 1's estimate of the distance at which orientation estimates break down is 23–32 arc min, at a viewing distance of 2.0 m. Jenkins (1983) estimate of the maximum dipole length facilitating a structure-vs-noise task is 1.4 deg measured at 57.3 cm. Scaling Jenkins' estimate (i.e., making the strong assumption that viewing distance will have little effect on performance) produces an estimate of 24.0 arc min: in agreement with the result of this experiment.

EXPERIMENT 2: EFFECT OF ORIENTATION VARIABILITY ON THE JUDGEMENT OF MEAN ORIENTATION

The aim of this experiment was to determine the effect of adding local orientation jitter on the judgement of the mean orientation of a Glass pattern. This form of noise maintains a constant distance between corresponding dots but adds uncertainty as to which *direction* a dot's correspondent lies in. This form of noise is important because one would expect that models relying on local orientation statistics will be critically affected by changes in these statistics.

The effect of adding small random rotations to dipole orientations was first observed by Glass & Switkes (1976), who informally demonstrated that such noise degraded perception of structure. This, they claimed, was consistent with the physiological model described by Glass (1969): the range of dipole directions now exceeds the specificity of a single orientation column and the excitation required to perceive structure is not achieved. Maloney *et al.* (1987) indicated that this range must be less than ± 11 deg because such a range of orientations does not significantly affect perception of structure.

Stimuli

Line and dipole textures similar to those used in the previous experiment were used, except that the length of elements was fixed at 8.0 arc min and the local orientation standard deviation (SD) of elements was systematically varied. Dipoles had Gaussian-distributed orientations with a SD of from 1.41 to 32.0 deg, sampled in multiplicative steps of $\sqrt{2}$. Examples are shown in Fig. 4.

Results

Threshold offsets for mean orientation judgement as a function of dipole orientation SD are shown for three subjects in the lower part of Fig. 4. Subjects typically achieve thresholds as low as 3.0 deg for the dipole textures and 2.0 deg for the line textures: an impressive level of performance given the large separation of dots and high density of dots. The pattern of results for the line and dipole textures are similar except that there is an approximately uniform shift (on logarithmic axes) of the functions from the lines to the dipoles. This indicates a multiplicative effect of matching uncertainty in the Glass patterns.

It is clear that there is little effect of adding orientation jitter on either the line or dipole textures until the SD exceeds about 8.0 deg. This figure is in accord with the value of ± 11.0 deg quoted in Maloney *et al.* (1987), which is equivalent to Gaussian-distributed orientations with a SD of 7.8 deg. Such a figure seems to indicate that there is inherent noise on the system which limits the accuracy of estimating mean orientation at very low levels of jitter. It is quite possible that this noise is due to the orientational bandwidth of the filters employed to extract structure. Beyond this level of orientation variation, performance deteriorates in an approximately power law relationship with orientation SD.

EXPERIMENT 3: EFFECT OF NUMBER OF ELEMENTS ON THE JUDGEMENT OF MEAN ORIENTATION

The aim of this experiment was to determine the effect that the number of elements making up a Glass pattern has on the accuracy of judging the mean orientation. A lack of effect of element number/density has been shown in the detection of bilateral symmetry in dot patterns (Jenkins, 1985), and in displacement limits for detection of motion in random binary luminance patterns (Morgan & Fahle, 1992) so one might expect Glass patterns to be similarly insensitive to density.

Stevens (1978) used patterns composed of dipoles located on a perturbed grid, which were not permitted to fall in such a way that alignment with nearby dipoles could cause "chains" of multiple elements. A rating judgement of "pairedness" was used to determine the maximum dipole separation for which structure was rated to be present. Dot density varied from 0.5 to 44 points/deg^2 (from 65 to 580 total dots) and the results indicated that, regardless of pattern type, if more than 2 or 3 points lay closer to a dot than its corresponding dot, then structure was not rated as present. Jenkins (1983), measuring the discriminability of Glass patterns from noise, produced results contradictory to Stevens', showing that there was no effect of altering dot density over the range 6.5 to 26.0 points/deg^2. Jenkins (1983) accounts for Steven's low estimate of tolerable noise by assuming that his subjects were conservative in their subjective rating of "pairedness". Another explanation for this inconsistency of findings is that it is the presence of low spatial frequency features like dipole "chains", which were eliminated from Stevens' stimuli, that indicates structure when dipole separation becomes large. Jenkins (1983) goes on to claim that stimulus field diameter is the critical factor in perceiving these patterns, independent of viewing distance.

Maloney *et al.* (1987) also measured the discriminability of Glass patterns from patterns of randomly oriented dipoles, but varied the number of unpaired noise dots added to the original pattern. Using two dot separations and a number of dot densities, they showed

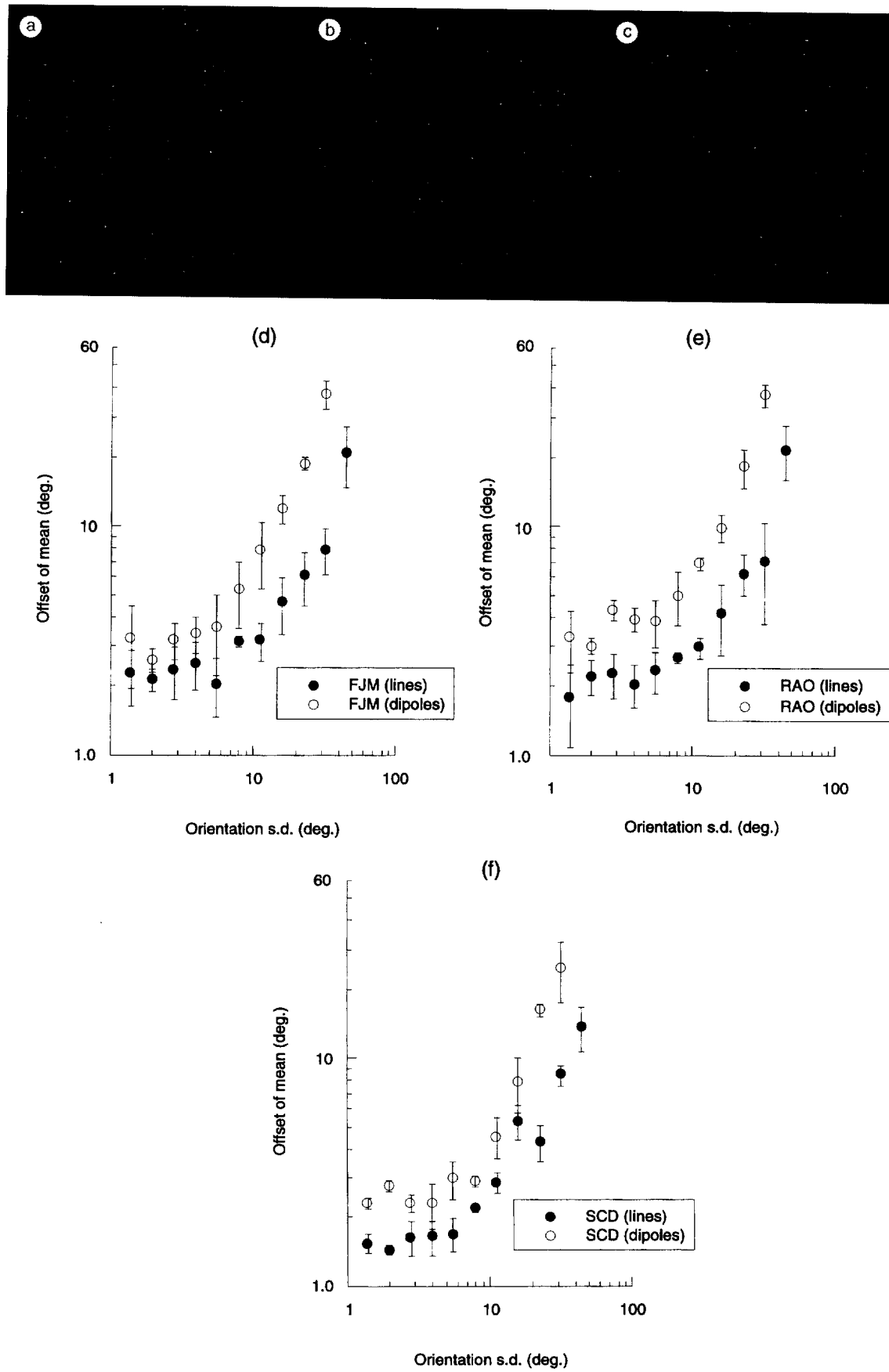
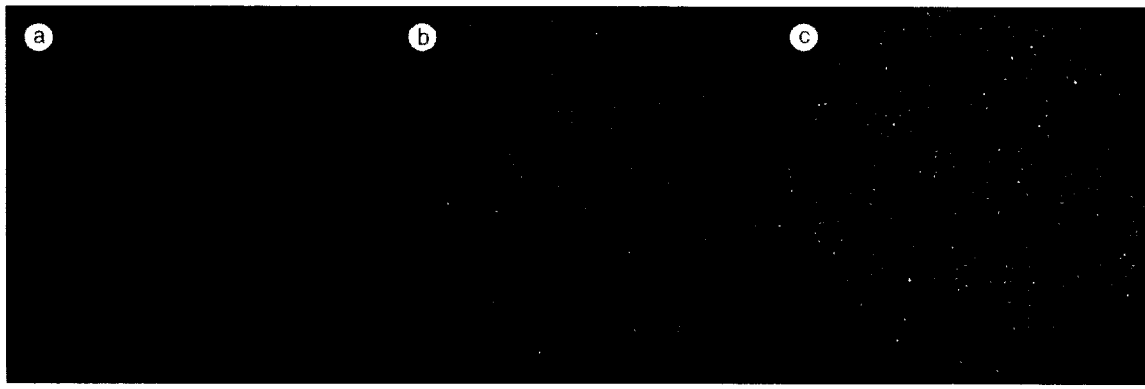
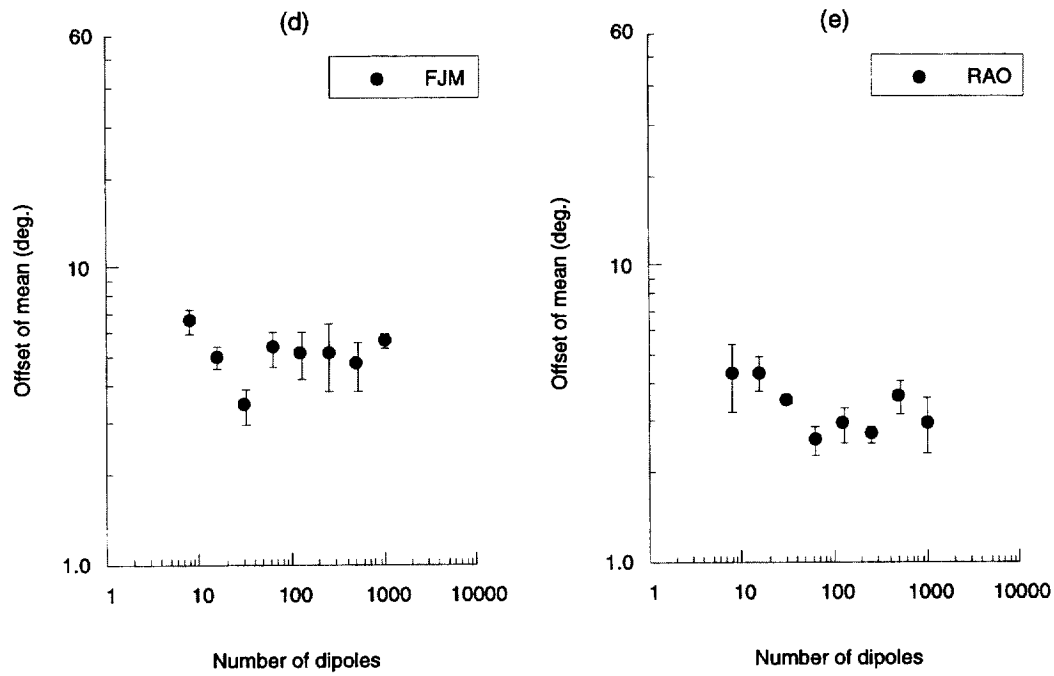


FIGURE 4. (a–c) Examples of the stimuli used in Experiment 2. The patterns have orientation variability with SD of (a) 1.0; (b) 4.0; and (c) 16.0 deg. (d–f) Threshold mean orientation offsets from three subjects as a function of element orientation variability.





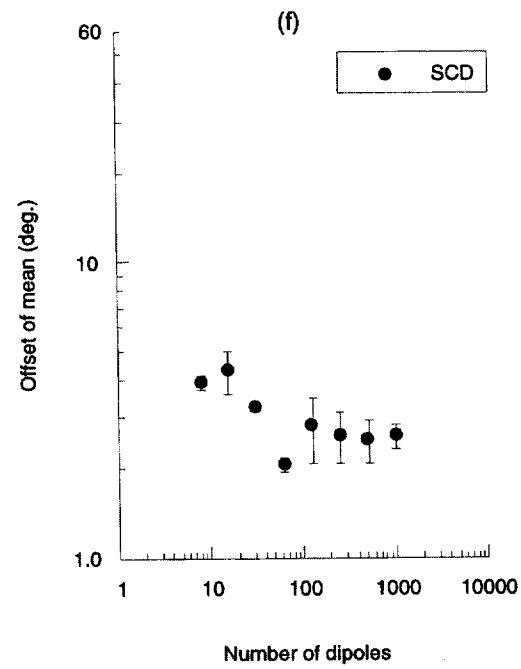


FIGURE 5. (a–c) Examples of the stimuli used in Experiment 3. The patterns shown contain (a) 64; (b) 256; and (c) 1024 dipoles. (d–f) Threshold mean orientation offsets of Glass patterns, as a function of the number of dipoles, from three subjects.

a Weber law relationship between the number of dipoles in the pattern and the maximum number of unstructured dots which could be tolerated at a particular level of performance. Their data show that detection at a level of d′=1.0 is possible with more than six noise dots closer to a dot than its partner. They suggest that this result argues against the neighbourhood approach of Stevens (1978) but, again, it probably indicates that comparison of these findings with Stevens' are precluded by differences between experimental procedures.

Pilot trials indicated that number/density had little effect with either the line or dipole textures, so this experiment was carried out only using Glass patterns.

Stimuli

All stimuli were Glass patterns with a dipole separation of 8 arc min, a value sufficiently large that any effects of neighbours should become apparent as density is increased. No dipole orientation variation was added. The number of dipoles in each pattern was varied from 8 to 1024 in one octave steps. Examples are shown in Fig. 5. Since constant field size and viewing distance were employed, dot density varied from 1.68 to 215 elements/deg^2.

Results

Threshold offsets for the mean orientation judgement as a function of the number of dipoles are shown for three subjects in Fig. 5. The accuracy of judging orientation is slightly poorer for very sparse patterns, but rapidly improves with increasing number of elements and performance asymptotes for patterns containing 32–64 elements. There also appears to be a slight dip in the function for all three subjects around 32–64 dipoles. The basic pattern of the data shows that there is little effect of stimulus density above about 64 dipoles. These data are in accord with performance on a structure detection task, reported in Jenkins (1983).

That subjects are relatively insensitive to pattern density does seem to be inconsistent with models based on neighbourhood matching. The performance of all three subjects is as good with patterns containing 1024 elements as with those containing 64. These patterns have, respectively, an average of 0.5 and 8.0 dots lying closer to each dot than its correspondent. This is in agreement with data from Maloney *et al.* (1987) which showed that structure vs no-structure judgements were possible when dots had more than six other dots closer to them than their correspondent.

TABLE 1. The stimulus parameters for Experiments 1–3

	Number	Length (arc min)	Orient. SD (deg)
Experiment 1	512	1.41–32.0	8.0
Experiment 2	512	8.0	1.41–32.0
Experiment 3	8–1024	8.0	8.0

SUMMARY—PSYCHOPHYSICS

To summarize the preceding experiments, the stimulus parameters are shown in Table 1.

The main results are as follows:

- Subjects are highly accurate at performing mean orientation judgements with thresholds which asymptote at around 1.5 deg for line textures, and 2.5 deg for Glass patterns.
- Estimating the mean orientation of a Glass pattern becomes easier as dipole separation is increased up to a critical separation, of around 5 arc min, beyond which performance rapidly deteriorates. For line textures there is a consistent improvement with increasing line length.
- Local orientation variation has little effect on judging the mean orientation of line textures or Glass patterns until a SD of around 8 deg is reached. A similar pattern of deterioration is observed with line and dipole textures, except that performance with Glass patterns is uniformly poorer.
- The accuracy of judging the mean orientation of a Glass pattern, within the limits tested, appears to be largely independent of the number of dipoles used.

MODELING OF MEAN ORIENTATION JUDGEMENTS

Four models for extracting features from texture are described: symbolic matching of tokens (Stevens, 1978), isotropic (Laplacian-of-Gaussian) filtering, oriented (Difference-of-Gaussian) filtering and "adaptive" oriented filtering. The performance of these models was compared with human data from the three conditions described. Versions of both the isotropic and adaptive filtering models incorporating a scale selection criterion, based on minimizing the SD of texture element orientation, are also described.

Isotropic filtering model

A model for texel extraction and mean orientation judgement was implemented using Laplacian-of-Gaussian filtering. It had five stages:

- Filtering with LoG at multiple spatial scales.
- Thresholding of filter responses.
- Symbolic description of resultant blobs.
- Calculation of mean blob orientation.
- Psychophysical decision.

The LoG is defined as:

$$\mathrm{LoG}(x, y, \sigma) = \left(1 - \frac{x^2 + y^2}{2\sigma^2}\right) e^{-(x^2+y^2)/2\sigma^2}$$

where σ refers to the space constant of the filter. The model was run using values of σ of 1.0 to $16\sqrt{2}$ pixels (0.57–12.9 arc min) in half-octave steps.

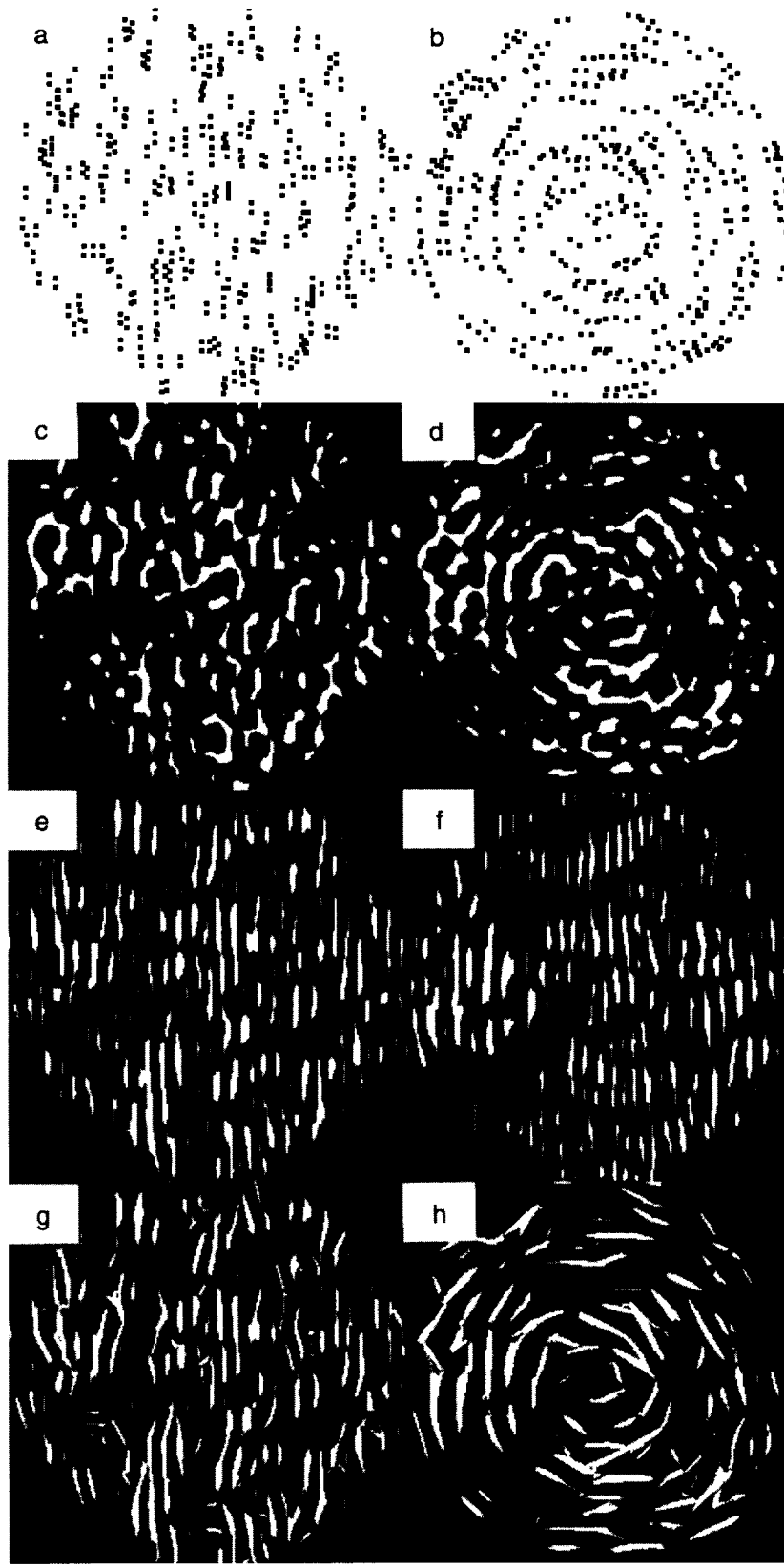


FIGURE 6. Operation of three pre-processing filtering schemes on two Glass patterns with (a) translational or (b) rotational structure. (c, d) LoG filtered and thresholded; (e, f) DoG filtered and thresholded; (g, h) adaptively filtered, at the same scale as (e, f), and thresholded.

Filter outputs were then doubly half-wave rectified, i.e., both positive and negative portions of the output were retained. The positive and negative blobs which result have previously been proposed as texture elements (Vorhees & Poggio, 1987, 1988), and more generally as the basic perceptual primitives of the raw primal sketch (Watt, 1988). The model described differs from the Vorhees & Poggio (1987) model in a number of minor ways. Firstly, this model has no gain control applied prior to filtering. Secondly, this model always thresholds at ± 1.0 SD of the filter output, rather than using histogrammed local intensity gradients. Generally the setting of the threshold is not critical; pilot simulations indicated that a threshold of anywhere between 0.75 and 2.0 SD is optimal for estimating local orientation. For this reason one would expect the predictions of this and Vorhees & Poggio's (1987) model to be broadly similar.

Zero-bounded regions in the thresholded image were then characterized using the image description scheme of Watt (1991). This produces compact "sentences" describing each blob in terms of its principal axis, centroid position, area, etc. Three features of each blob were used to calculate the mean orientation and the orientation variability of the set: orientation, mass and aspect ratio. The blob aspect ratio, and mass were used as a measure of the *reliability* of the orientation of each blob; large, elongated blobs are more reliable than small, near-circular blobs. Calculation of the mean orientation and orientation variability at each spatial scale is described in the Appendix.

Finally, the psychophysical decision was made by classifying orientations between 0 and 90 deg as clockwise of vertical, and those between 90 and 180 deg as anticlockwise of vertical.

The model described so far makes a decision based on information present at any one spatial scale. A version of the model was implemented which automatically selected filter size by minimizing blob orientation variability.

Oriented filtering models

Motivated by the presence of cells in V1 which are not only sensitive to the spatial scale of a pattern, but also to its orientation, a number of models have been proposed for deriving local orientation estimates using oriented filtering. Zucker (1982) has proposed a model which estimates image "flow" direction using the identity of the most locally active oriented Difference-of-Gaussian filter, in conjunction with a relaxation algorithm which maximizes orientation consistency (i.e., co-linearity) within a neighbourhood.

The two models described in this section also use DoGs but differ fundamentally from the model described in Zucker (1982). Firstly, they use not the identity of a filter to estimate orientation but the filter output (and subsequently a symbolic description) from which orientation estimates are made. Secondly these models do not incorporate iterative post filtering. The first uses convolution with individual DoG filters, the second ensures local orientation consistency by picking the peak filter output, on a point-by-point basis, at the convolution stage.

The first model operates on a DoG filter centred on the orientation which patterns varied around from trial to trial (i.e., vertical), thereby assuming prior knowledge of the pattern orientation. The two-dimensional (2D) DoG is composed of a DoG in the x-direction multiplied by a Gaussian in the y-direction:

$$W(x_t, y_t) = (e^{-x_1^2/2\sigma^2} - (1/2.23)e^{-x_1^2/2(2.23\sigma)^2})e^{-y_t^2/2(3\sigma)^2}$$

where σ refers to the SD of the positive Gaussian function. x_t and y_t are co-ordinates rotated by angle ϕ:

$$x_t = x\cos\phi + y\sin\phi$$

$$y_t = y\cos\phi - x\sin\phi$$

The ratio of the amplitudes of the positive and negative parts of the DoG and the aspect ratio are based on those derived by Wilson and co-workers using a variety of psychophysical paradigms (Phillips & Wilson, 1983; Wilson & Gelb, 1984). A range of filter sizes was employed with σ varying from 1.0 to $8\sqrt{2}$ pixels (0.57–6.46 arc min) in half-octave steps. Examples illustrating convolution of a DoG filter with a Glass pattern are given in Fig. 6.

The problem with using the output of a single DoG filter is how to deal with images that contain more complex orientation structure. How are filter outputs integrated across orientation? Inspired by the presence of intra-orientation inhibition between cells with similar orientation selectivity (Morrone & Burr, 1986; Morrone *et al.*, 1982; Tsumoto *et al.*, 1979), the second oriented filtering model uses point-by-point selection of the most active DoG filter. This was assessed by computing the squared filter output across 12 orientations, at each spatial scale, and adding a small amount of Gaussian blur to each of these local energy representations. At each point in the image, local energy is compared across orientation and the (unsquared) output of the filter with greatest energy selected. This is a simpler version of Malik & Perona's (1990) "leaders-take-all" system. It may also be considered a type of "steerable" filter as described by Freeman & Adelson (1991). However, in order to avoid the restrictions on the aspect ratio of a mechanism derived by summation of a small set of basis filters (Freeman & Adelson, 1991) the steering is performed explicitly on the output of oriented filters. This is computationally inefficient but apparently unavoidable if one is to model the known characteristics (i.e., aspect ratios around 1:3) of orientationally tuned mechanisms in human vision. Examples of the operation of this mechanism, here termed "adaptive" filtering, are shown in Fig. 6(g, h). Note that the model produces highly oriented blobs compared with the LoG [Fig. 6(c, d)] although they have more complex shape than the single DoG output [Fig. 6(e, f)].

Beyond the initial filtering stage, the details of both models are identical to the LoG model. A symbolic description of thresholded blobs is constructed, mean

orientation estimated, and the psychophysical decision made.

Token matching algorithm

A token matching model described in Stevens (1978) was implemented. The model calculates all possible pairings of a single dot to other dots in a surrounding neighbourhood, and all possible pairings of those dots in similar sized regions around them. All matches have a corresponding orientation by which they are histogrammed (weighted by the proximity of the matched dot to the original dot). Smoothing of the local histogram is performed by using a relatively small numbers of "buckets", and the peak orientation is then selected. Since correctly paired dots will form dipoles which are locally parallel, the peak orientation should match the orientation of the correct pairing. So the corresponding dot is determined by selecting the virtual line whose orientation most closely matches the peak orientation. If no line can be found within 15 deg of the peak, no solution is returned for the dot. So far, this is a direct implementation of the model described in Stevens (1978). Having derived a set of virtual lines and their orientations, the model is extended to estimate the mean and makes a decision as to whether the texture is clockwise or anticlockwise.

Proximity weighting in Stevens' model is relative to the neighbourhood size. The weighting of a virtual line's contribution to a local orientation histogram is either 1, 2/3 or 1/3 depending on whether neighbouring dots are less than 1/4, less than 1/2 or greater than 1/2 a neighbourhood radius apart, respectively.

In the simulations described, two methods of setting the neighbourhood size were examined. The first set the radius equal to the dot separation, i.e., the optimal size for discounting unmatched dots. Stevens (1978) claims that, since subjects cannot see structure when more than two or three dots lie closer to a dot than its correspondent, such a small region will not tend to give enough samples to allow the reliable extraction of a peak orientation. Stevens claims that a neighbourhood which contains six or seven dots closely emulates human performance on the psychophysical tasks he describes. This was tested in the simulation which follows by setting the size of a neighbourhood so that it would contain, on average, 6.5 dots.

Simulation method

To compare model predictions to psychophysical data, simulations of Experiments 1–3 were performed. In order to generate a mean orientation threshold a method of constant stimuli was used. Sixty-four stimuli were generated at each of 17 stimulus levels (which were adaptively sampled in the psychophysical experiments). Each stimulus image was processed using one of the four models described, an estimate of the mean orientation extracted, and the model's response recorded. The SD of the resultant psychometric function was then calculated.

Because of processing time constraints, the scale of analysis was chosen, for models incorporating a scale selection criterion, using the first six stimuli at the beginning of a run and that filter size employed during the rest of the run. Thus, each time that the independent variable was changed (e.g., dipole length) the choice of scale was made again by running the full model over the first six stimuli and using the mean spatial scale that the criterion specified. The neighbourhood size parameter of the Stevens' model was set by hand at the beginning of each run.

Simulation results: isotropic filter model

Figure 7(a) shows the SD of the mean orientation estimate as a function of dipole length for individual LoG filters. It is clear that for progressively larger dipole lengths, coarser scale filters give the best estimates of mean orientation. It is also apparent that no response from any one filter can explain the variation in subjects' performance on this task. The solid line shows predictions of a model incorporating the scale selection criterion. Note that the overall pattern of responses is broadly similar to subjects. The primary difference is that the best performance of the model occurs around a narrow range of dipole separations about 2.5 arc min, whereas human performance is best around a broader range (2.5–5 arc min) of separations.

Accuracy of the LoG model compared with human subjects, as a function of additional orientation jitter, is shown in Fig. 7(b). Again the model produces the same pattern of responses as human observers, but this time does not approach their best performance on this task at low levels of orientation SD. Human subjects consistently achieve thresholds of around 3.0 deg, compared with the model whose best performance is around 5.0 deg. Although this is a small difference it is important because it suggests that a model based on the Laplacian-of-Gaussian *cannot explain the basic level of performance in this task.* Deviations of a model from data which are due to the model exceeding human performance can be explained in terms of noise on the system. This type of deviation cannot. Figure 7(b) also shows that the critical level of orientation SD, beyond which performance deteriorates, for the model (around 20.0 deg) is greater than for human observers (around 8.0 deg).

It is possible that the failure of the model on this condition is due to the setting of certain parameters of the model, such as the degree of thresholding, the use of a single spatial scale, etc. To try and at least partially take this into account, simulations of Experiment 2 were rerun, with three different levels of thresholding (0.5, 1.0 and 2.0 grey-level SDs), and incorporating four different levels of integration across scale (average across blob orientations from ± 1 or ± 2 octaves of spatial scale). These variations did not produce any improvement in performance beyond that presented in Fig. 7(b).

While this does not preclude the possibility that some other treatment of the LoG output might produce better results, it does at least suggest that the result is not an

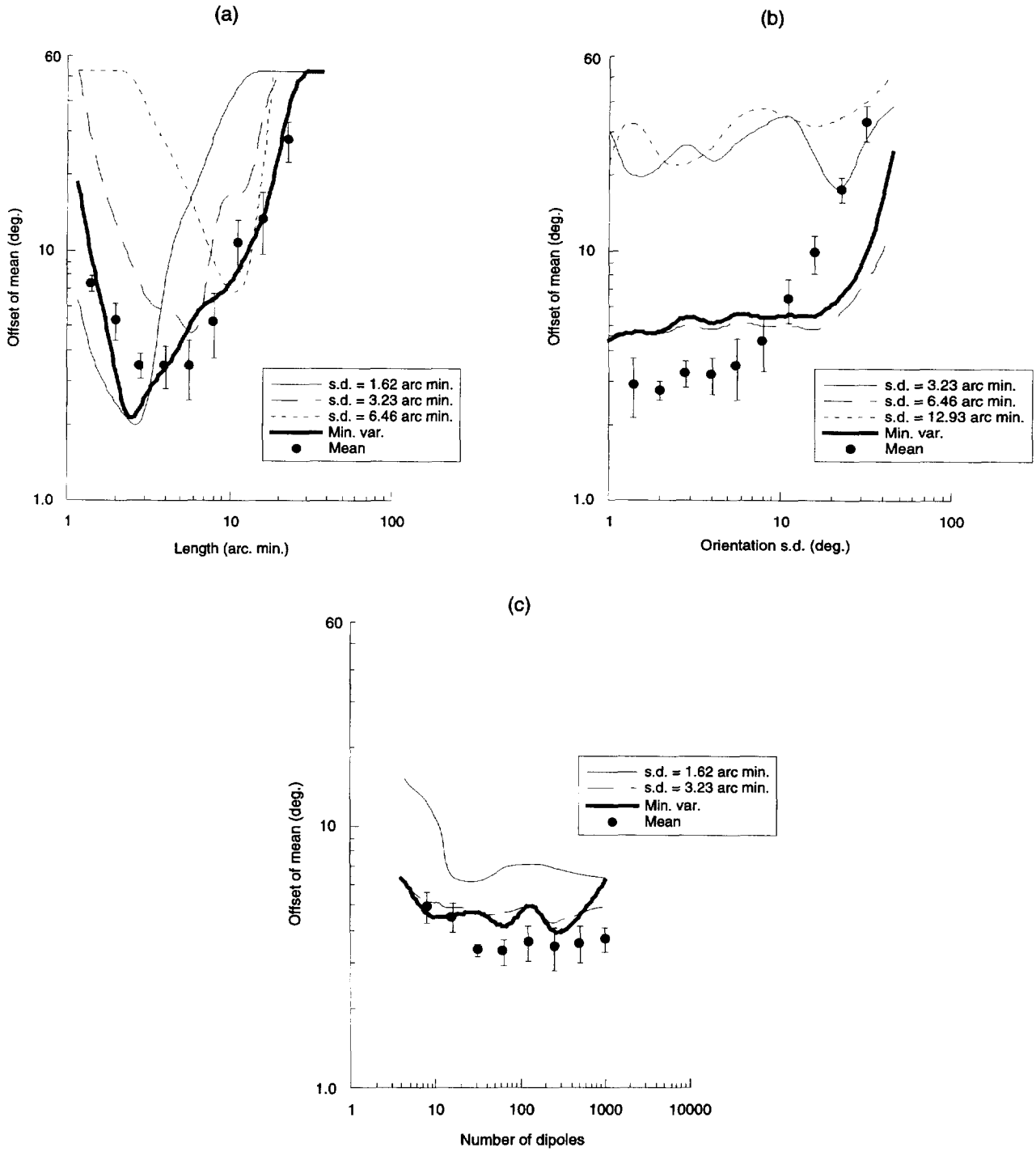


FIGURE 7. Threshold offset for LoG estimates of mean orientation of dipole textures, as a function of (a) dipole length; (b) dipole orientation SD; and (c) dipole number. Thin lines are the predictions from LoGs operating at a single scale, thick lines the prediction of a filter operating at the scale minimizing element orientation SD.

artifact of the setting of some individual variable within the model as described.

Finally, Fig. 7(c) compares predictions from the LoG model and data from the density condition, Experiment 3. There is a reasonable match between human data and the predictions from the model incorporating automatic scale selection, although the basic level of performance of the model is again slightly worse than data.

In summary, this section has suggested that a model which uses LoG filtering, to extract and describe texture primitives, shows a similar pattern of results to the data from Experiments 1 and 3. However, this model fails to explain the performance of subjects in Experiment 2.

Simulation results: oriented filter models

The result of the simulation of the mean orientation judgement as a function of dipole length, using the two oriented filter models, is shown in Fig. 8(a). Filled

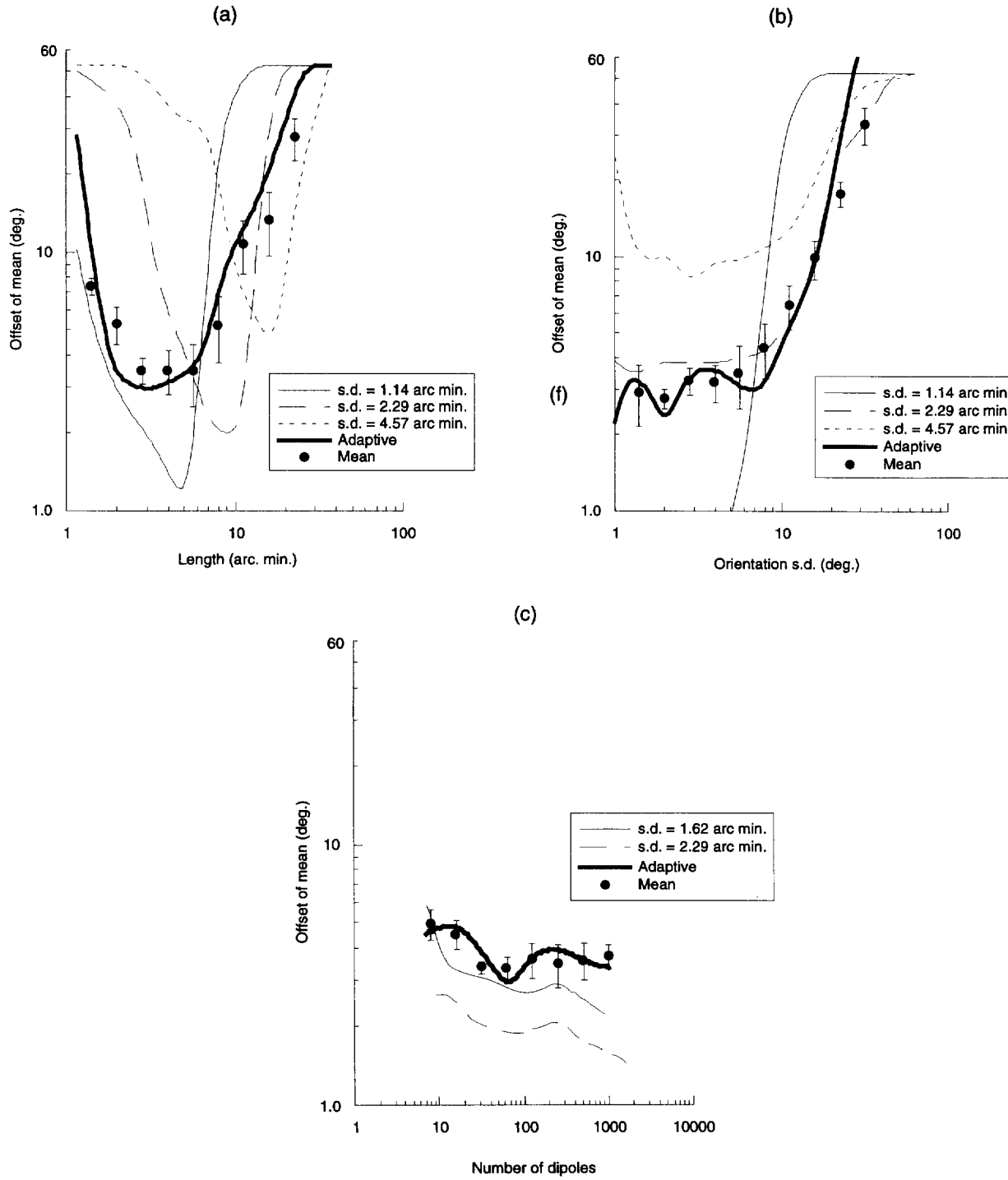


FIGURE 8. Threshold offset for oriented filter estimates of mean orientation of dipole textures, as a function of (a) dipole length; (b) dipole orientation SD; and (c) dipole density. Thin lines are the predictions from a model with a single orientation/scale of filter, the thick line represents the predictions of the adaptive filtering model operating at the scale minimizing texel orientation standard deviation.

symbols represent the mean performance of the three subjects from Experiment 1, fine lines the predictions of individual DoG filters, and the coarse line the prediction of the adaptive filtering model. It is clear that *no one DoG filter can explain subjects' performance on this task.* If filter size is too small, or large, compared with the separation of the dipoles, only uniformly poor estimates can be made of mean orientation. The performance of these filters declines suddenly as the length of dipoles exceeds the size of the excitatory zone of the oriented filter; there is a small range of lengths for which a particular filter is optimally tuned.

Predictions from the adaptive filtering model, operating at a scale determined by minimizing texel orientation variation, are uniformly poorer than the single DoG models but closely match the performance of subjects.

Note that no fitting has been applied to the model predictions.

The mean performance of the three subjects from Experiment 2 is shown in Fig. 8(b) along with predictions from the single DoG and adaptive filtering models. Results suggest that the output of single DoG filters show the same pattern of deterioration in estimates of mean orientation, as a function of local orientation SD, as the human subjects. Furthermore, the output of a single DoG with σ between 4.90 and 9.8 arc min fits subjects' performance well. Predictions from the adaptive model, shown as the heavy line, again match subjects' data well. Because dipole length was constant, the adaptive filtering model behaves very much like a single DoG model in this condition.

Figure 8(c) shows the simulation results for the task from Experiment 3. Results are similar to those from the last simulation: a single DoG filter with a SD between 4.90 and 9.8 arc min shows the same trend as human data, as does the adaptive filtering model.

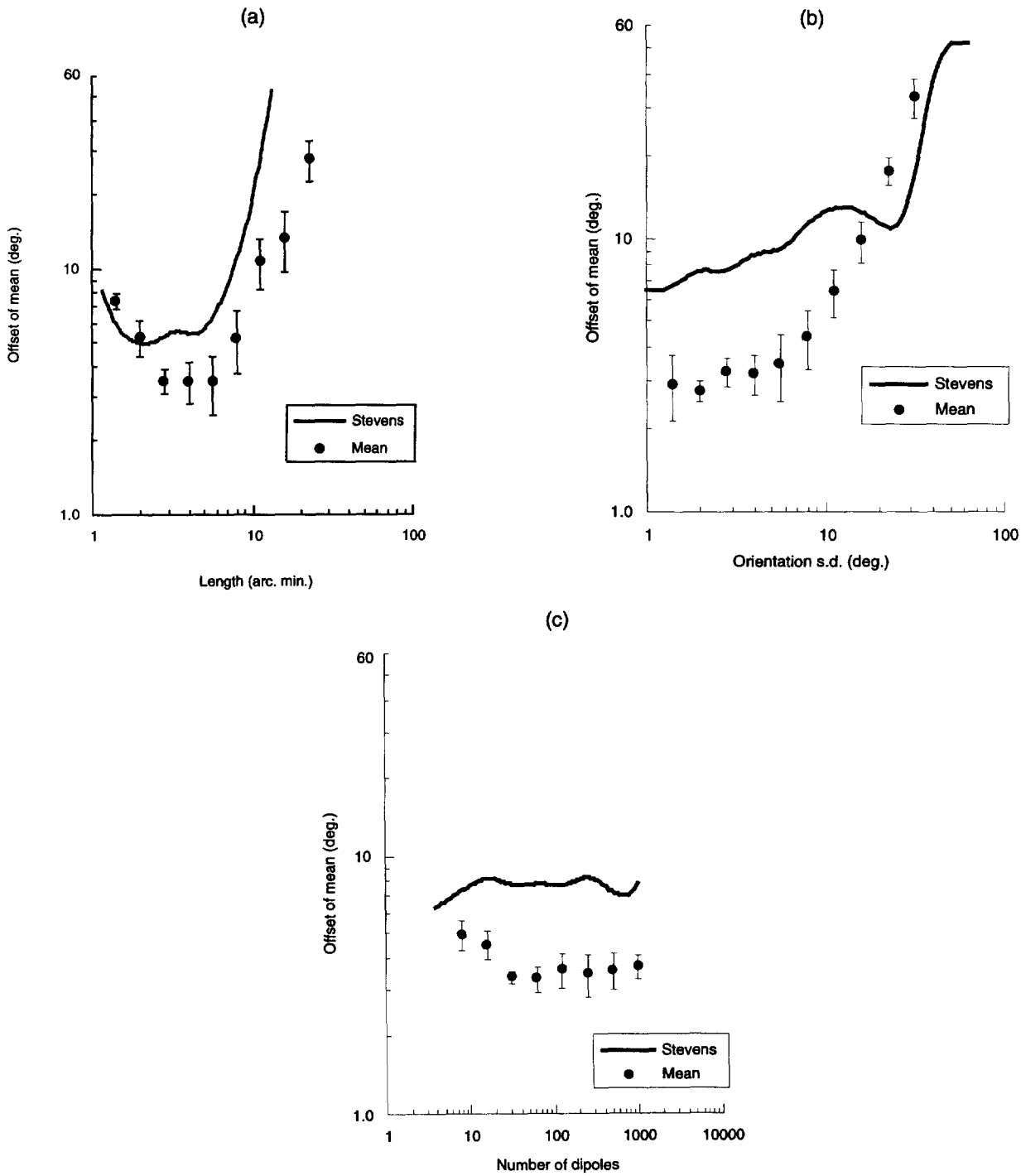


FIGURE 9. Threshold offset for token matching model estimates of mean orientation of dipole textures, as a function of (a) dipole length; (b) dipole orientation SD; and (c) dipole density.

Simulation results: token matching model

Predictions from the token matching model on the three tasks are compared with human data in Fig. 9. Even though the predictions shown are for the model operating using the optimal neighbourhood size for the pattern, they are uniformly worse than both human subjects and the other models described.

Predictions from the model using the pattern density to set neighbourhood size are uniformly poor and are not presented. Given the way this model operates, this strongly suggests that in detecting structure in Glass patterns subjects make use of the *low spatial frequency information* which arises from accidental co-alignments of dipoles. Stevens' model can only use *individual* dipole orientations and when dipole separation is large in relation to dot density, as was the case in our experiments, matching of individual dipoles breaks down. It is concluded that this model can provide an explanation for the perception of structure in Glass patterns in only the most limited of cases.

SUMMARY—SIMULATIONS

Weighted chi-squares of the fits of the three models to the subjects' data, from the three tasks using dipole textures, are given in Table 2. The token matching model is rejected in all conditions. The LoG model is not rejected for the orientation variability and dipole density conditions but is for the length condition. This failure, along with the failures of Stevens' model, are important because both models' performance in these cases was *worse* than that of human subjects. The adaptive filtering model is not rejected in any condition, its chi-square values being (ordered by condition) 55, 27 and 7% of values associated with the next-best fit. On the grounds of parsimony it is accepted as the best model of subjects' performance in the experiments.

GENERAL DISCUSSION

Grouping stimuli, such as Glass patterns, allow the study of feature extraction from texture because they isolate orientation as a useful source of information at a narrow range of spatial scales. Judgement of the mean orientation of these patterns as a function of dipole length, orientation jitter, and number suggests that the visual system accurately selects the correct filter size which gives the best estimate of mean orientation. An appropriate spatial scale of analysis minimizes the orientation SD of resultant features. When combined with an estimate of local orientation, measured using a form of oriented filtering, this criterion proved to account adequately for the data from the three conditions. From the experiments reported we conclude that, in order to derive Glass pattern structure with sufficient accuracy, spatial filtering accounts of feature extraction from texture are constrained in two ways. Firstly, an isotropic, Laplacian-of-Gaussian mechanism does not suffice. Secondly, because Glass patterns are spatially broad-band stimuli containing structure only at a narrow range of spatial scales, the output of a narrow range of filters must be available. How do these findings fit in with established models of visual processing?

TABLE 2. Weighted chi-squares of the fits of the three models to the subjects' data from the three tasks using dipole textures

	d.f.	Auto Adaptive	Auto LoG	Token matching
Experiment 1	8	0.90*	1.65*	7.71
Experiment 2	9	1.31*	4.87	5.31
Experiment 3	7	0.15*	2.02*	4.83

*Goodness-of-fit measure fails to reject the model at the 0.05 level.

Oriented rather than isotropic mechanisms

Our sensitivity to the orientation structure of Glass patterns exceeds the predictions of at least one class of isotropic filter. This is clearly problematic for models using LoGs to derive features (Vilnrotter *et al.*, 1986; Vorhees & Poggio, 1987, 1988; Wen & Fryer, 1991) or more generally for construction of the "primal sketch" (e.g. Marr, 1976, 1982), but is consistent with the presence of orientationally selective channels in human vision (e.g. Hubel & Wiesel, 1967). However, because isotropic filtering schemes require only one convolution per spatial scale they are not only an *efficient* way of deriving local features but have also avoided the question of how the outputs of channels at different orientations are combined. The failure of the LoG model discounts at least one combination rule: linear summation (which would equate to isotropic DoG filtering), although there may be circumstances in which this rule does hold (e.g., for the judgement of appearance, Georgeson, 1992). The results presented here suggest that the combination rule is probably nonlinear allowing the output of more active channels to dominate. This is necessary if one is to retain the orientational resolution provided by oriented filters. The nonlinear combination rule (peak selection) used in the adaptive filtering scheme is one possible rule. Others, such as the "leader-takes-all" scheme of Malik & Perona (1990), are equally plausible.

Selection of spatial scale

It is a general, and largely unaddressed, problem of computational visual processing to select the appropriate spatial scale of analysis for a task. Scale-space filtering (Witkin, 1983), proposes that features such as zero-crossings (ZCs) are "tracked" through spatial scale and their persistence used to judge their utility. However, this approach has not been expanded from one- to two-dimensional signals. The scale space approach constructs tree-like maps of ZCs and relies on the fact that as one proceeds to lower frequencies features can only ever merge and new ones never appear. While this is proven for Gaussian-filtered one-dimensional signals (Babaud *et al.*, 1986), it has been shown that is not the case for two dimensions (Yuille & Poggio, 1986). Thus, a more complex representation is required.

It almost certainly over-ambitious to propose that

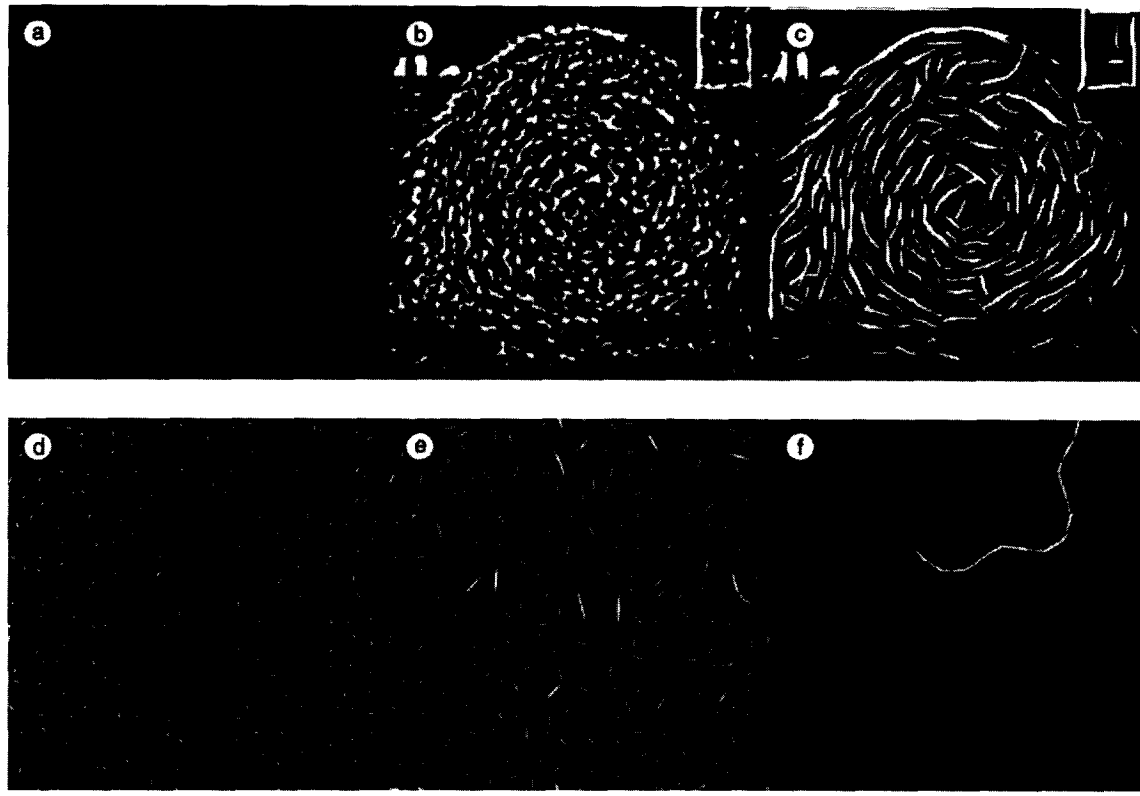


FIGURE 10. Grouping properties of the adaptive filtering model. (a) Haystack image and blobs derived using (b) LoG and (c) adaptive filters. Note that the features derived in (c) are much more oriented than those in (b). (d) Typical stimuli from Field *et al.* (1993). A "path", with successive elements differing in orientation by ± 45 deg, is embedded in a field of randomly oriented elements. (e) The adaptively filtered version of (d) showing aggregation of the elements in the patch. (f) By generating a symbolic description of (e), the path may be derived automatically by isolating the longest feature.

information at a range of scales is collapsed into a single description that will be useful for everything. I suggest that "task constraints" critically influence spatial scale selection for visual processing. The scale selection criterion described (minimization of local orientation variability) is a useful one for deriving reliable estimates of local orientation. However, the visual system undoubtedly has a number of ways of determining the appropriate spatial scale depending on what is to be done with the information derived. For example, Elder & Zucker (1996) have proposed the use of the "minimum reliable scale" for detecting edges, using local estimates of the likelihood of error due to sensor noise. It may be that these various criteria for maximizing reliability of information are all fundamentally statistical. Human visual processing of texture disregards precise spatial localization of elements, and concentrates on representing trends in populations of features (e.g. texture edges). It may be that a major role of texture processing is the estimation of image statistics for scale selection.

Texture perception and contour integration

Figure 10(a) shows an oriented texture filtered with (b) LoG and (c) adaptive filters with similar peak spatial frequency sensitivities. Because the adaptive filtering model integrates in the direction of local contour orientation it produces accurate estimates of local orientation from such natural images. The output of the LoG is much more sensitive to local noise. Clearly the processes involved in the extraction of tokens from texture are implicated in contour integration, an area of increasing interest within psychophysics (e.g. Field *et al.*, 1993; Hess & Field, 1995) The results reported here suggest that measurement of local contour orientation must be made using oriented filters. More specifically, Fig. 10(e) demonstrates that the path integration process itself might be achieved by the adaptive filtering mechanism described here. Contours within the stimuli used by Field *et al.* (1993) are readily grouped by the adaptive filtering process. Furthermore by constructing a symbolic description and simply selecting the longest feature in the image, the path may be automatically extracted [Fig. 10(f)]. Note that the adaptive filtering model relies only on local excitation of similarly oriented filters. No inhibition between orientations that are mutually inconsistent with the presence of a contour [a component of the "association field" model proposed by Field *et al.*, (1993)] is necessary.

REFERENCES

Appelle, S. (1970). Perception and discrimination as a function of stimulus orientation: the "oblique" effect in man and animals. *Psychological Bulletin, 78*, 266–278.

Babaud, J., Witkin, A. & Duda, R. (1986). Uniqueness of the Gaussian kernel for scale-space filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 8*, 26–33.

Brodatz, P. (1966). *Textures: a photographic album for artists and designers*. Toronto, Canada: Dover.

Caelli, T. & Julesz, B. (1979). Psychophysical evidence for global feature processing in visual texture discrimination. *Journal of the Optical Society of America, A69*, 675–678.

Elder, J. H., & Zucker, S. W. (1996). Local scale control for edge detection and blur estimation. In *Lecture Notes in Computer Science: Proceedings of the 4th ECCV '96*, (Vol. 1064, pp. 57–69). New York: Springer.

Field, D. J., Hayes, A. & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local "association field". *Vision Research, 33*, 173–193.

Freeman, W. H. & Adelson, E. H. (1991). The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13*, 891–906.

Georgeson, M. A. (1992). Human vision combines oriented filters to compute edges. *Proceedings of the Royal Society of London B, 249*, 235–245.

Glass, L. (1969). Moiré effects from random dots. *Nature, 243*, 578–580.

Glass, L. & Perez, R. (1973). Perception of random dot interference patterns. *Nature, 246*, 360–362.

Glass, L. & Switkes, E. (1976). Pattern recognition in humans: correlations which cannot be perceived. *Perception, 5*, 67–72.

Hess, R. F. & Field, D. J. (1995). Contour integration across depth. *Vision Research, 35*, 1699–1711.

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

Jenkins, B. (1983). Spatial limits to the detection of transpositional symmetry in dynamic dot textures. *Journal of Experimental Psychology: Human Perception & Performance, 9*, 258–269.

Jenkins, B. (1985). Orientational anisotropies in the human visual system. *Perception and Psychophysics, 37*, 125–134.

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.

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

Maloney, R., Mitchison, G. & Barlow, H. (1987). Limit to the detection of Glass patterns in the presence of noise. *Journal of the Optical Society of America A, 4*, 2336–2341.

Marr, D. (1976). Early processing of visual information. *Proceedings of the Royal Society of London B, 275*, 483–534.

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

Morgan, M. J. & Fahle, M. (1992). Effect of pattern density upon displacement limits for motion detection in random binary luminance patterns. *Proceedings of the Royal Society of London, 248*, 189–198.

Morrone, M. & Burr, D. (1986). Evidence for the existence and development of visual inhibition in humans. *Nature, 321*, 235–237.

Morrone, M., Burr, D. & Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells. *Proceedings of the Royal Society of London B, 216*, 335–354.

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

Prazdny, K. (1986). Psychophysical and computational studies of random-dot Moiré patterns. *Spatial Vision, 1*, 231–242.

Stevens, K. (1978). Computation of locally parallel structure. *Biological Cybernetics, 6*, 19–28.

Tsumoto, T., Eckart, W. & Creutzfeldt, O. (1979). Modification of orientation sensitivity of cat visual cortex neurons by removal of GABA-mediated inhibition. *Experimental Brain Research, 34*, 351–363.

Vassilev, A., Simeonova, B. & Zlatkova, M. (1981). Orientation acuity at detection threshold. In *Proceedings of the Fourth European Conference on Visual Perception*, (p. 17).

Vilnrotter, H., Nevatia, R. & Price, K. (1986). Structure analysis of natural textures. *IEEE Transactions on Pattern Analysis, and Machine Intelligence, PAMI-8*, 679–698.

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).

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

Wagemans, J., Van Gool, L., Swinnen, V. & Van Horebeek, J. (1993). Higher-order structure in regularity detection. *Vision Research, 33*, 1067–1088.

Watt, R. J. (1988). *Visual processing: computational, psychophysical and cognitive research*. London: Lawrence Erlbaum.

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

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

Wen, W. & Fryer, R. (1991). Texture boundary detection—a structural approach. In Mowforth, P. (Ed.), *Proceedings of the British machine vision conference* (pp. 104–110).

Westheimer, G. (1981). Visual hyperacuity. *Progress in Sensory Physiology, 1*, 1–20.

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. P. (1983). Scale-space filtering. In *Proceedings of the 8th International joint conference on artificial intelligence* (pp. 1019–1021). Karlsruhe, West Germany.

Yuille, A. L. & Poggio, T. (1986). Scaling theorems for zero-crossings. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 8*, 15–25.

Zucker, S. (1982). *Early orientation selection and grouping: evidence for type I and type II processes*. Montreal, Canada: McGill University.

Acknowledgements—I would like to thank Roger Watt and Ian Paterson for many helpful discussions and their co-development of much of the software used in this study. Thanks are also due to Robert Hess and Isabelle Mareschal for their comments on earlier drafts of this paper. This work was supported by the SERC (grant GR/H53181) while the author was at the University of Stirling.

APPENDIX

Given a set of n estimates of orientation $\theta = \{\{\theta_k\}: 1 \leq k \leq n, 0 \leq \theta \leq \pi\}$, one cannot use the arithmetic mean of the angle values to represent the sets' overall orientation, because orientation is a cyclical dimension (angles can differ by a maximum of 90 deg). The mean of a data set minimizes the difference between all members of the data set and itself. The arithmetic mean assumes that subtraction measures this difference; this is not the case for two angles.

Assuming a non-uniform distribution of data (i.e., one for which the mean is defined), a set of measures has a mean orientation ($\bar{\theta}$). A measure of deviation from the mean is:

$$e_n = \left\{ \begin{pmatrix} |\theta_n - \bar{\theta}| & \text{if } |\theta_n - \bar{\theta}| < \frac{\pi}{2} \\ \left(\frac{\pi}{2} - |\theta_n - \bar{\theta}|\right) & \text{otherwise} \end{pmatrix} \right.$$

By the principles of least squares, the mean of the data set should minimize the quantity:

$$E = \sum_{k=1}^{n} e_k^2 \tag{1}$$

Because E is discontinuous, an analytic minimization is not possible but it is straightforward to minimize this quantity iteratively to an arbitrary level of precision. Given an estimate of the mean orientation, E provides the measure of orientation variability used for the model described above. Because of computational considerations, a differentiable

alternative for e is also considered: the square of the vector product. Assuming all vectors are unit length, E is approximated by:

$$E = \sum_{k=1}^{n} \sin^2(\theta_k - \bar{\theta}) \tag{2}$$

and the maximum/minimum of this is to be found where:

$$\frac{dE}{d\theta} = \sum_{k=1}^{n} \sin 2(\theta_k - \bar{\theta}) = 0.$$

It follows that:

$$\sum_{k=1}^{n} \sin 2(\theta_k \cos 2\bar{\theta}) = \sum_{k=1}^{n} \cos 2(\theta_k \sin 2\bar{\theta})$$

and hence:

$$\bar{\theta} = \frac{1}{2} \tan^{-1} \left[\frac{\sum_{k=1}^{n} \sin 2\theta_k}{\sum_{k=1}^{n} \cos 2\theta_k} \right].$$

This expression gives a value which is guaranteed to yield an extremum which may be a maximum or a minimum. In the latter case $\bar{\theta}$ will be 90 deg greater than the true mean orientation. To resolve the ambiguity one must evaluate the second derivative which, for a maximum, must be less than zero:

$$\frac{d^2E}{d\theta^2} = -2 \sum_{k=1}^{n} \cos 2(\theta_k - \bar{\theta}) < 0.$$

So the final expression for deriving the mean orientation is:

$$\bar{\theta} = \left\{ \begin{pmatrix} \bar{\theta} & \text{if } \frac{d^2E}{d\theta^2} < 0 \\ \bar{\theta} + \frac{\pi}{2} & \text{otherwise} \end{pmatrix} \right.$$

The orientation variance is calculated by inserting the estimated mean into Eq. (2).