

Orientation Variance as a Quantifier of Structure in Texture

STEVEN C. DAKIN

5 McGill Vision Research Unit, Department of Ophthalmology, 687 Pine Avenue West, H4-14,
Montreal, Quebec, Canada H3A 1A1.
Email: scdakin@vision.mcgill.ca
Fax: (CAN) 514 843 1691

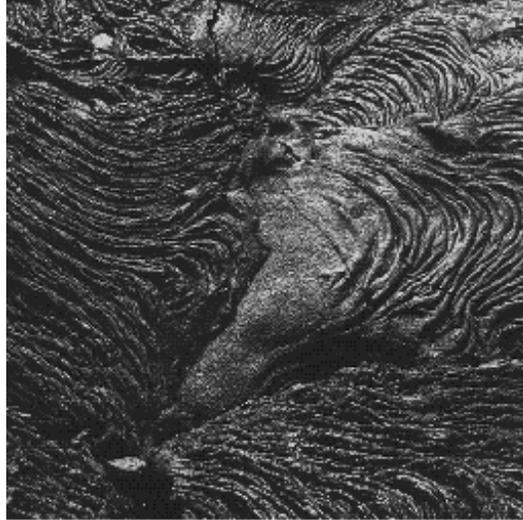
10

Abstract - I consider how structure is derived from texture containing changes in orientation over space, and propose that *multi-local orientation variance* (the average orientation variance across a series of discrete images locales) is an estimate of the degree of organization that is useful both for spatial scale selection and for discriminating structure from noise. The oriented textures used in this paper are *Glass patterns*, which contain structure at a narrow range of scales. The effect of adding noise to Glass patterns, on a structure versus noise task (Maloney, Mitchison & Barlow, 1987), is compared to discrimination based on orientation variance and *template matching* (i.e. having prior knowledge of the target's orientation structure). At all but very low densities, the variance model accounts well for human data. Next, both models' estimates of tolerable orientation variance are shown to be broadly consistent with human discrimination of texture from noise. However, neither model can account for subjects' lower tolerance to noise for translational patterns than other (e.g. rotational) patterns. Finally, to investigate how well these structural measures preserve *local orientation discontinuities*, I show that the presence of a patch of unstructured dots embedded in a Glass pattern produces a change in multi-local orientation variance that is sufficient to account for human detection (Hel Or & Zucker, 1989). Together, these data suggest that simple orientation statistics could drive a range of "texture tasks", although the dependency of noise resistance on the pattern type (rotation, translation, etc.) remains to be accounted for.

15

20

25



30

FIGURE 1. Image of lava flow illustrating complex orientation structure.

1. INTRODUCTION

1.1 Constraints on the representation of orientation

35 Because the physical processes underlying the formation of texture are subject to complex interactions, the orientation structure of surfaces in natural scenes is often complex. Consider the image of lava flow shown in Figure 1. The patterns of ripples in the rock are due to differences in the speed of the magma flow at the time of formation. Many possible factors contribute to these differences (irregularities in the underlying surface, etc.) and their effects are recorded in such complex flow patterns. Notice that, although the orientation structure is complex,
40 orientation within small areas is relatively uni-directional. This suggests a constraint on what is meant by a “globally organized” orientation field: a set of regions may be defined such that a single orientation is representative of the orientation of features within each region. More formally we would expect that the *local orientation variance* of texture elements within each region will be low for a smooth orientation field. We refer to the average of those variance
45 estimates as the *multi-local orientation variance*.

In attempting to derive structure from such images, the visual system’s spatial representation of local orientation is constrained in two ways. First, it must assign a representative orientation to each of a number of regions. This reduces the redundancy and complexity of the representation and allows for the effects of noise. This constraint maintains
50 *smooth changes of local orientation in space*. The presence of *abrupt changes* in local orientation, e.g. at the edges of the smooth region in the centre of Figure 1, should also be maintained. Such *texture boundaries*, which have been very intensively studied psychophysically (for review see Bergen, 1991)¹, in analogy with luminance, are thought to signal surface discontinuities. How true this is for natural images has yet to be demonstrated.

55 To summarize: spatial averaging of local orientation that occurs during visual processing is constrained to maximize confidence in estimated local orientation while retaining interesting discontinuities. Satisfying both constraints significantly restricts the scale of integration. Too fine a scale will result in vulnerability to noise, but too coarse a scale will “smooth out” orientation discontinuities. Multi-local orientation variance is a simple measure of the
60 smoothness of the orientation field.

¹ This is due to the influence of Julesz (e.g. Julesz, 1981), and Marr, who proposed a map of local surface discontinuities, based on various visual cues, as the basis of the 2.5D sketch (Marr, 1976; Marr, 1982).

1.2 Glass Patterns

Evidence for the utility of local orientation statistics in texture processing comes from studies of Glass patterns (Glass, 1969). These textures are composed of a set of randomly positioned dots with a geometrical transformation of the same dot set superimposed upon it. The resulting stimulus yields a striking percept of global structure based on the matching of dots to their transformed counterparts (Figure 2d). Grouping of dots with their correspondents (collectively known as *dipoles*) can be achieved using spatial filters (e.g. Zucker 1982; Dakin, 1997a,b). However in Glass patterns, spatial frequency bands adjacent to informative frequencies contain high-energy noise making this matching process critically dependent on filter size (Prazdny, 1986; Dakin 1997a). The judgment of the mean orientation of translational Glass patterns is well predicted using “adaptively” selected oriented filters whose spatial scale is set by minimising the orientation variance of features in the filter-output (Dakin, 1997a).

For translational textures orientation variance may be computed from the entire pattern. The goal of this paper is to define simple *local* orientation statistics and to test their general usefulness as a measurement of local “structure” in Glass patterns both computationally and psychophysically. Experiments 1 and 2 consider the utility of a multi-local orientation variance measure, both as a scale selection criterion and as a discriminator of structure-from-noise. Discrimination based on orientation variance is compared to an intuitively appealing alternative: the *template matching* model. The latter has *a priori* knowledge of the structure of the target and makes an explicit comparison of local orientation estimates to the expected orientation at each position. In Experiment 3, I investigate how well these models capture the disruption to orientation structure caused by the presence of orientation discontinuities. Specifically the problem of detecting anomalous patches embedded in oriented textures is considered. A modification of the paradigm presented by Hel Or & Zucker (1989) is presented, and data are simulated using multi-local orientation variance and template models.

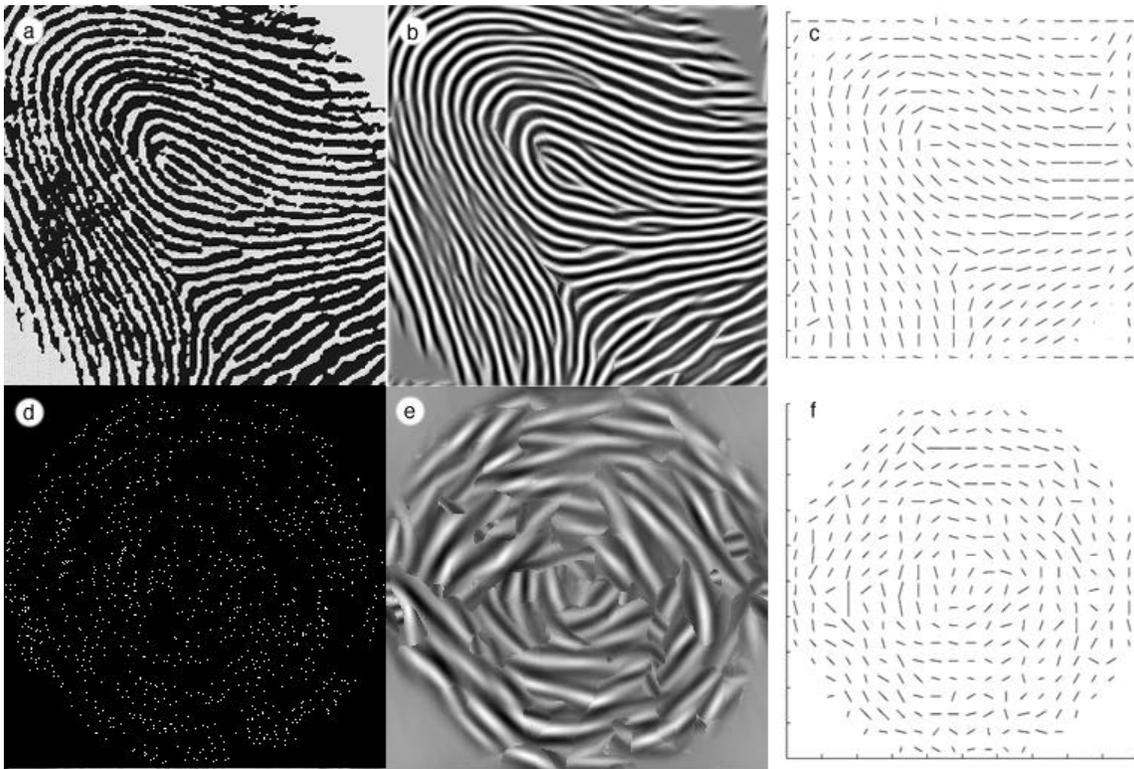


FIGURE 2. Estimating local orientation using adaptive filtering. (a) Fingerprint image, and (d) Glass pattern; both images contain complex orientation structure. (b,e) Adaptive filter output, (c,f) flow fields derived using the scheme described in Appendix 1.

2. MEASURING THE DEGREE OF STRUCTURE

For simple translational textures, I have previously used a filter-based procedure to estimate the global smoothness of the orientation field (Dakin, 1997a). This model consists of:

1. “Adaptive” filtering at multiple spatial scales: convolution with a bank of oriented filters followed by pixel-wise selection of the output of the most active filter across orientation.
2. Extraction of texture elements/features. At each spatial scale, the combined filter outputs are thresholded and a symbolic description of resulting “blob” attributes (orientation, position, etc.) is computed.
3. Computation of multi-local orientation variance, and selection of the scale which minimizes this value.

Section 1 of the Appendix contains a summary of Stages 1 and 2 of the model (described in more detail in Dakin, 1997a) which extract oriented features (or texture elements) from the image. Because this paper considers texture that contains complex orientation structure, Section 2 describes a scheme for partitioning the image and Section 3 extends Stage 3 of the model to produce three simple *multi-local* orientation statistics: mean orientation, orientation variance, and deviation of local orientation from an expected value. Figure 2 illustrates the operation of the extended model for calculation of local mean orientation from two textures. Note the highly elongated structure in the output of the adaptive filtering stage (shown at one spatial scale, Figure 2 b,e), and that the flow-fields derived from the output of such filters reflect local orientation (Figure 2c, f).

3. EXPERIMENT 1: EFFECT OF UNCORRELATED DOTS ON STRUCTURE DETECTION

This experiment was designed to apply the extended model to globally organized Glass patterns, and to compare its discrimination of structure from noise, to that of human observers. Maloney *et al.* (1987) present the first attempt to quantify the strength of the Glass pattern percept (Glass, 1969) using signal detection theory. Subjects were presented with a pattern that was either pure noise or a Glass pattern with added randomly positioned dots. They were then required to make a forced-choice classification of the stimulus as structured or not. Data from Maloney *et al.* (1987), using a dilation of 18 arc min. and a rotation of 9 arc min., are plotted in Figures 3 and 4, respectively. Discriminability (d') decreases with increasing noise for all dot densities tested. Furthermore, there is a Weber law dependence of d' on the ratio of paired to unpaired dots (i.e. d' is proportional to the ratio on a log-linear plot). In this section we consider whether the model described for quantifying local orientation structure predicts such a dependence.

3.1 Details of simulation

The adaptive filtering model was used to isolate texture elements (the “blobs” illustrated in Figure 2e) from input patterns. Multi-local orientation variance was computed for each spatial frequency and the model operated at the spatial frequency which minimised this value. Given a description of local orientation, two strategies were tested for deciding whether a pattern was an organized or a random texture (both are described in detail in Section 3 of the Appendix). The first, multi-local orientation variance, is the same estimate used to select spatial scale, and is expected to be lower for organized than for random patterns. The second strategy tested was a *template match*: total difference between local estimates of orientation and the expected orientation at that position, given *a priori* knowledge of the underlying orientation transformation.

Simulations used stimuli similar to those from Maloney *et al.* (1987): rotations of 9 arc min. and translations of 18 arc min. For a pattern containing D dipoles and N noise dots, simulations were run at $D=10, 25, 50, 100$ and 200 . and $N=D, 2D, 2D, 2D...16 2D$. At each stimulus level, 32 organized and random patterns were generated, and an estimate of the magnitude of the

variance and template-matched cues was calculated for each pattern. The 32 values for each pattern were stored and used to calculate pattern discriminability, d' , at each stimulus level.

145 Given two sets of estimates of organisation, from the noise and texture images at a given signal to noise ratio, the criterion (c) for deciding if a pattern was noise or texture was set at the point that divides the means of the two sets of estimates (μ_1 and μ_2) in the ratio of the standard deviations of each set (σ_1 and σ_2):

$$c = \mu_1 + \frac{\sigma_1}{\sigma_1 + \sigma_2} (\mu_1 + \mu_2) \quad (1)$$

150

One thousand random selections were made from the sets of organized and random cue and each compared to this criterion. If the cue from an organized pattern exceeded criterion this trial was classified as a hit. If the cue from a random pattern exceeded criterion, this trial was classified as a false alarm. Then d' was calculated from hr , the fraction of structured patterns labeled as structured and fa , the fraction of random patterns labeled as structured with:

155

$$d' = P^{-1}(hr) + P^{-1}(fa) \quad (2)$$

where $P^{-1}(y)$ is the inverse function of the Gaussian probability function:

160

$$P(y) = \frac{1}{\sqrt{2\pi}} \int_0^y e^{-x^2/2} dx \quad (3)$$

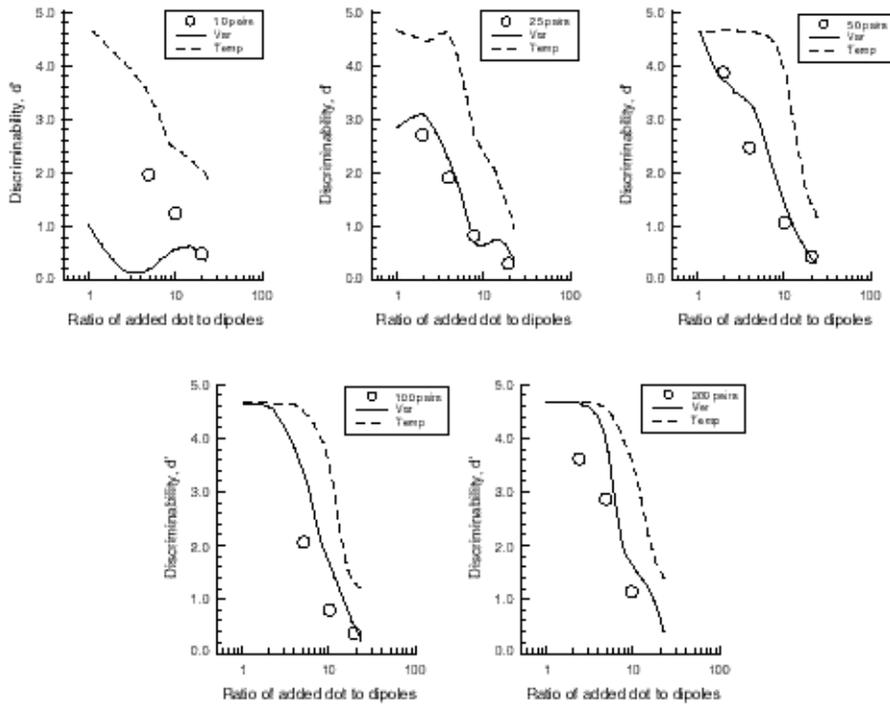
Values of this function can be calculated to an arbitrary level of precision using an approximation to the incomplete Gamma function (given in Press, Teukolsky, Vetterling & Flannery, 1992). However, if no false alarms occur, or if the hit rate is 100%, then the value of d' is meaningless and only related to the number of trials. To allow the assignment of some value to d' in these cases, the maximum hit rate was fixed at 99% and the minimum false alarm rate fixed at 1%. In this way the maximum possible d' was 4.653, which is the maximum tabulated value in Swets (1964) (the source used by Maloney *et al.* (1987) for calculating d').

170

| | | 10 els. | 25 els. | 50 els. | 100 els. | 200 els. |
|-----------------|-------------|---------|---------|---------|----------|----------|
| Variance | Translation | 0.701* | 0.784 | 0.213* | 0.0242* | 0.220 |
| | Rotation | 1.77 | 0.319* | 0.325* | 0.979 | 0.371* |
| Template | Translation | 0.532* | 0.306* | 0.451 | 0.824 | 0.298* |
| | Rotation | 5.42 | 12.4 | 9.37 | 9.96 | 4.3 |

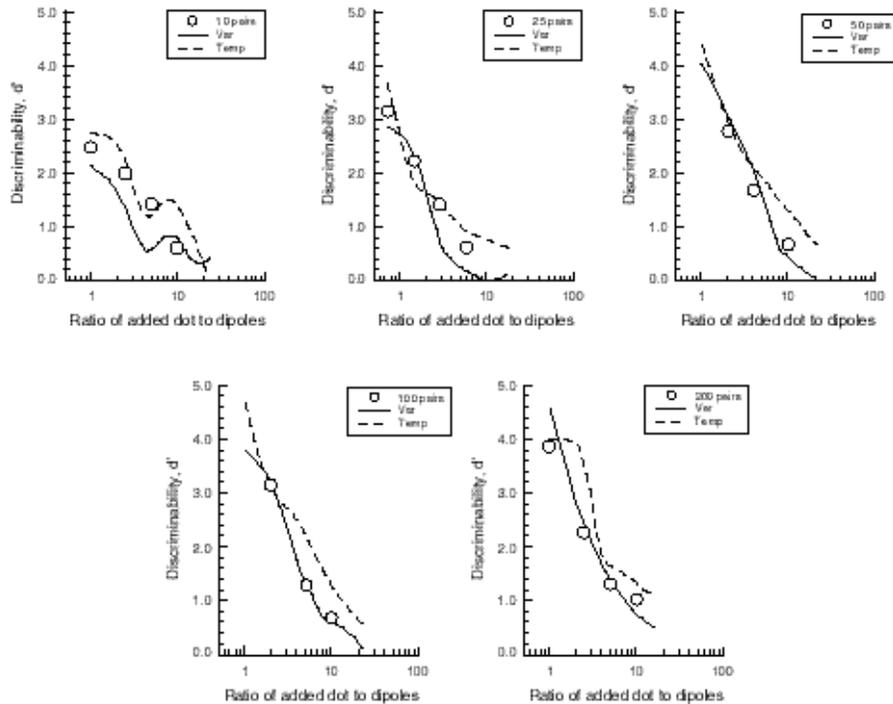
TABLE 1. Chi-squares for the fits of the variance and template models to the data from Maloney, *et al* (1987). ("els" = elements").

175



180

FIGURE 3. Discriminability of 9 arc min. rotational Glass patterns from random pattern by human observers (unfilled circles) and two models. The lines labeled "MMB" show predictions from the ideal observer model of Maloney *et al.* (1987). The other lines, marked "Var" and "Temp", show predictions from a field smoothness measure and a template match, respectively, operating on the output of the adaptive filtering model.



185 FIGURE 4. As Figure 3, for 18 arc min. translational Glass patterns.

3.2 Simulation results

Results from the simulations are shown in Figure 3 (for rotational fields) and Figure 4 (for translational fields). The template model performs at least as well as human subjects in the ten conditions shown. However, as the total number of dipoles increases this model displays a pronounced plateau around low signal-to-noise ratios (SNRs) indicating perfect discrimination. Note that neither model shows a pure linear dependence on the ratio of added dots to paired dots - as predicted by Maloney *et al.* (1987). Both models plateau for high SNRs (i.e. highest ratios of organised to noise dots). Although no such plateau is evident in data from Maloney *et al.* (1987), performance at the lowest ratio tested (1:1) often approaches $d' = 4.0$ (indicating near perfect discrimination), and therefore at higher SNRs performance must plateau.

The discrepancy between the template model and human data is greatest for the task using rotational fields, shown in Figure 3. This transformation used a dipole length of 9 arc min. compared to 18 arc min. for the translational field. This suggests that the template model is able to make the most of low uncertainty on *local* orientations to produce near perfect discrimination. The variance model on the other hand is, proportionally, less disrupted by uncertainty on local orientation. This makes sense given that the variance operator combines multiple orientation measures, giving resistance to local orientation uncertainty but losing information in the process. Both the variance and template models show similar slopes to the data for virtually all densities tested. The variance model produces closer fits to human performance than the template model in most conditions (see Table 1), usually due to the very high performance of the template model. The agreement of both models and data is impressive given that no free parameters were used to fit predictions. However the variance model fails to achieve human levels of performance on very sparse rotations (Figure 3a) and translations (Figure 4a). There are a number of possible explanations for this failure. First, the smoothing used to calculate the orientation field might be insufficient, leading to “holes” in the orientation field and a deflation in the estimate of field smoothness. Second, the field might be too sparse for a multi-local orientation variance measure to be useful. In the limit, the template model can operate on just one orientation estimate as, given *a priori* knowledge of what

215 transformation to expect, can the subject. The variance measure requires multiple orientation estimates, which sparse patterns may not provide.

In conclusion, while the superior accuracy of the template matching measure is clear, the multi-local variance model provides a generally good account of the results reported by Maloney *et al.* (1987) and correctly predicts an approximately Weber law relationship between additive noise and number of dipoles, for the range of signal to noise ratios tested. The template model performs impressively in all conditions, producing perfect detection at SNRs up to 1:8, but consequently often overestimating human performance. It may be that there is an attentional component to the template matching model that limits its utility at higher densities and leads subjects to rely more heavily on local orientation statistics. A prediction arising from this view is that a subject's *uncertainty* about the transformation underlying a texture should be most damaged at low feature densities.

4. EXPERIMENT 2: EFFECT OF ORIENTATION VARIANCE ON STRUCTURE DETECTION

The model described uses multi-local orientation variance to determine the spatial scale and it would therefore seem sensible to compare its performance to that of human subjects for textures with variability imposed on local orientation. In the following experiment the effect of element length, and the addition of local orientation variance on the perception of structure in Glass patterns, were again examined using a "structure versus noise" paradigm. Specifically the maximum amount of tolerable orientation variance was determined using an adaptive method, as a function of element length. This method has been applied elsewhere using mesh-derived textures (Hallet, 1992), and vector patterns (Dodwell & Caelli, 1985).

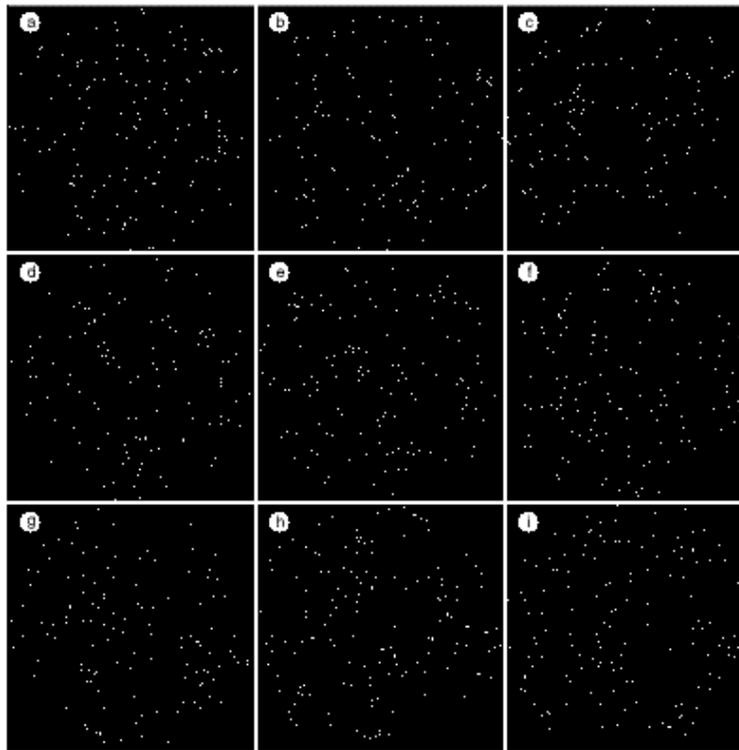
235 4.1 Method

4.1.1 Subjects

Five subjects took part in the experiment. SCD, DW and SW had corrected-to-normal vision. Subjects other than the author were naïve to the purpose of the experiment and relatively inexperienced with psychophysical procedures. A short training period (64 presentations) was undertaken before data collection began.

4.1.2 Apparatus

An Apple Macintosh IIfx computer was used to generate and display stimuli on a Formac ProNitron 80.21 colour monitor (frame refresh rate of 75Hz). Subjects viewed the screen binocularly, with natural pupils at a distance of 2.0 metres. All stimuli were presented at the centre of the screen, indicated by a pre-stimulus fixation marker. Stimulus duration was 100ms followed by a 750ms ISI.



250 FIGURE 5. Examples of the stimuli used in Experiment 2. Each row shows three Glass patterns composed of 256 dipoles of the same length: (a-c) 4 pixels, (d-f) 8 pixels, (g-i) 16 pixels. Images in each row differ in the amount of orientation variation added to dipoles: (a, d, g) $\pm 0^\circ$, (b, e, h) $\pm 22.5^\circ$, (c, f, i) $\pm 45^\circ$.

4.1.3 Stimuli

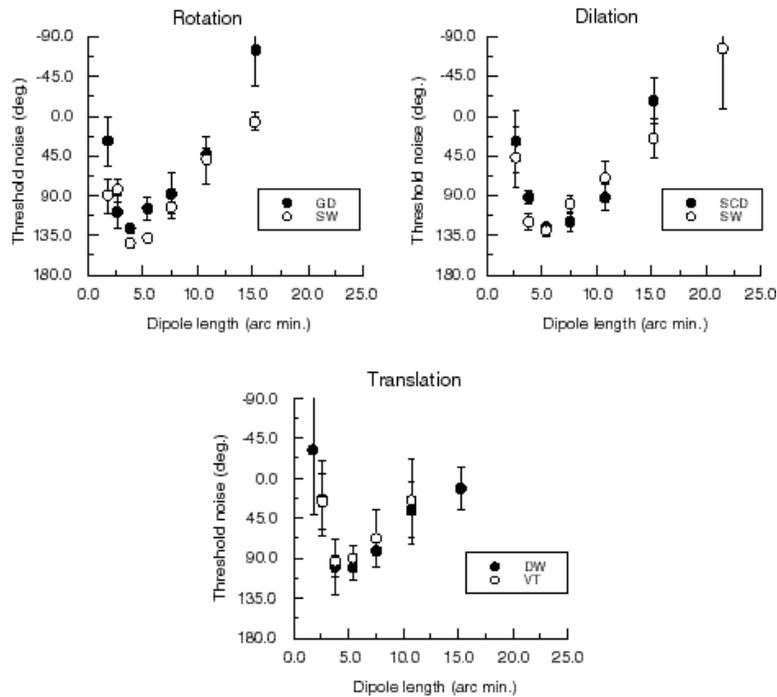
The stimuli used were approximately circular texture fields (radii subtended 1.23 degrees) composed of dot pairs (dipoles). Dipoles were randomly distributed in the field and appeared white on a black background. Fields appeared within a 2.46 degree (256 pixel) square image. 255 The initial orientation of dipoles was set according to one of three transformations: a rotation, a dilation or a vertical translation (examples are shown in Figure 5). Note that only *orientation* was determined by the transformation; if dipoles were the result of applying a true rotation or dilation upon a random set of dots then their length would depend on their distance from the pattern's focus which would have provided an additional cue to the presence of the Glass pattern. 260 While a similar length transformation to the random dipole pattern could have compensated for this, it is orientation which is the primary attribute of interest in this experiment and for this reason only the orientation components of the rotation and dilation transformations were used.

265 This orientation derived from the particular transformation was then subjected to a uniform random shift, the size of which was systematically varied. A range of dipole separations was tested from 2-23 arc min. Conditions were not intermixed within a run.

4.1.4 Procedure

270 The subjects' task was a two interval, 2AFC. In one randomly selected interval the texture was a Glass pattern, in the other a texture composed of randomly oriented dipoles. The subject indicated which of the two textures was a Glass pattern by depressing one of two keys on the computer keyboard. The independent variable was the amount of orientation jitter placed on the orientations of the dipoles in the Glass pattern. APE, an adaptive method of constant stimuli (Watt & Andrews, 1981) sampled a range of jitters from $\pm 90^\circ$ (where discrimination from the noise texture is at chance) to $\pm 0^\circ$ in steps of 22.5° . The threshold orientation jitter was

275 estimated from the slope of the resulting psychometric function and is equivalent to the 83% correct point. If subjects are unable to reliably discriminate structure from noise in the absence of orientation disruption, the threshold noise level is effectively interpolated from the measurable portion of the psychometric function and, for that reason, can be negative.



280 **FIGURE 6.** Results from three subjects for Experiment 2. Graphs show tolerable jitter of dipole orientation, for
 285 a structure versus noise discrimination, as a function of dipole length. Orientation jitter is the width (in
 degrees) of a uniform random orientation disruption imposed on dipole orientation. 180° on the ordinate is
 equivalent to random orientations, 0° a perfectly aligned orientation field. The U-shaped function indicates
 there is a narrow range of dipole lengths which produce optimal resistance to noise for this task.

4.2 Results

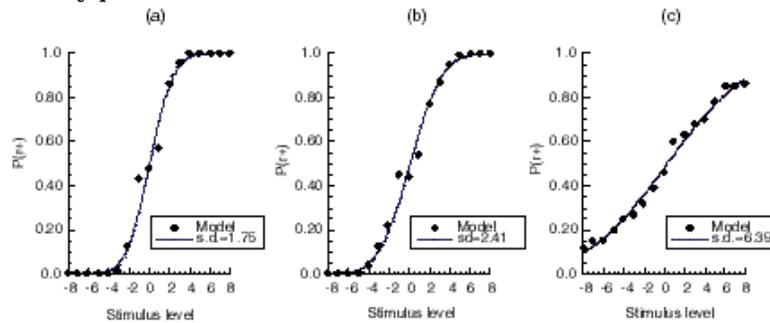
290 Results with two subjects for three transformation types are shown in Figure 6. Data show a U-
 shape trend: for all transformation types tested there are a small range of dipole separations at
 which the subject can withstand the most orientation jitter, typically at dipole lengths of 4-8
 arc min. At this separation, subjects can reliably discriminate between noise and Glass patterns
 which are composed of dipoles which have local orientations jittered by $\pm 56^\circ$.

295 A decrease in tolerable orientation variance as length increases is to be expected. As
 matching uncertainty increases the effect of orientation variability will be more disruptive.
 The poor performance exhibited with textures composed of very short elements could be due to a
 number of factors. Even though there is little uncertainty as to the matching of the components
 of dot pairs, subjects are poorer at estimating the orientation of features with small aspect
 ratios (Westheimer, 1981). This is confirmed by data from Experiment 1 of Dakin (1997a)
 which showed that subjects are optimal at estimating the mean orientation of translational
 300 Glass patters at dipole separations of around 4-5 arc min. One complication, however, is in the
 use of the threshold orientation variance as a measure of performance. Because dots are
 sampled from a pixel array, variance might be expected to have an unpredictable effect on
 performance at very low separations. Pixelation effectively quantises the orientations that can
 be displayed so that the effect of variability could be amplified as orientations are forced into
 the next "bin". A final possible reason for reduced performance at small dipole separations, is
 305 that the spatial frequency of mechanisms sensitive to such dipoles is very high. Poor
 performance may be attributable to dipoles failing to exceed the contrast threshold of such
 filters.

310 A second feature of the data is the poorer resistance to noise for translational Glass patterns compared to either of the globally organized patterns. t-tests indicate significant differences between data for the translational case and data from both rotational and dilational patterns ($p < 0.05$, 6 d.f.). Differences between performance with rotational and dilational patterns is not significantly different. Similarly, Wilson, Wilkinson & Asaad (1997) have reported that translational Glass patterns are less resistant to the intrusion of noise (in the form of added uncorrelated dots) than rotational patterns. They interpret this result as an *advantage* for rotational patterns implicating concentric structure detectors (identified with the operation of cells in V4).
315

4.3 Details of simulation

320 The two measures of pattern structure, field smoothness and template match (described above) were compared to the task described. Simulations were run using identical stimuli to the psychophysical experiment; the same range of dipole lengths and the same transformations were simulated. However, the full version of the adaptive filtering model uses 7 spatial scales and 12 filter orientations. Even using fast Fourier techniques it was not feasible to simulate 50 psychometric functions using 96 separate convolutions for every image. To avoid this problem a simplified version of the model was used. At the beginning of each new run, at each dipole length, the average optimal spatial scale was calculated over 16 Glass patterns and subsequent convolutions were only performed at this scale.
325



330 FIGURE 7. Typical psychometric functions from the simulation. These figures show the performance of the variance discriminator for dipole separations of 4.0, 8.0 and 16.0 arc min.

4.4 Simulation results

335 Psychometric functions were calculated at each dipole length condition by generating a set of 32 organized textures at each of the nine stimulus levels $\pm 0^\circ$ to $\pm 90^\circ$ in steps of 22.5° as used in the psychophysical experiment. The variance and template-matched estimates were calculated for the pattern and these 32 values for the patterns were stored. A similar procedure was used to generate structure estimates for a reference/disorganized Glass pattern. The psychometric function could then be generated by sampling a pair of values from a file derived from cued stimuli and from one derived from the reference stimuli. The proportion of times one was greater than the other represents the probability that the model would discriminate a stimulus with a particular cue from the reference. Examples of psychometric functions derived in this way are given in Figure 7. Note that they are well fitted by the standard cumulative Gaussian model:
340

$$P(r) = \frac{1}{\sqrt{2}} \int_{-\infty}^r \exp \left[-\frac{(e - \mu)^2}{2} \right] de \quad (4)$$

345

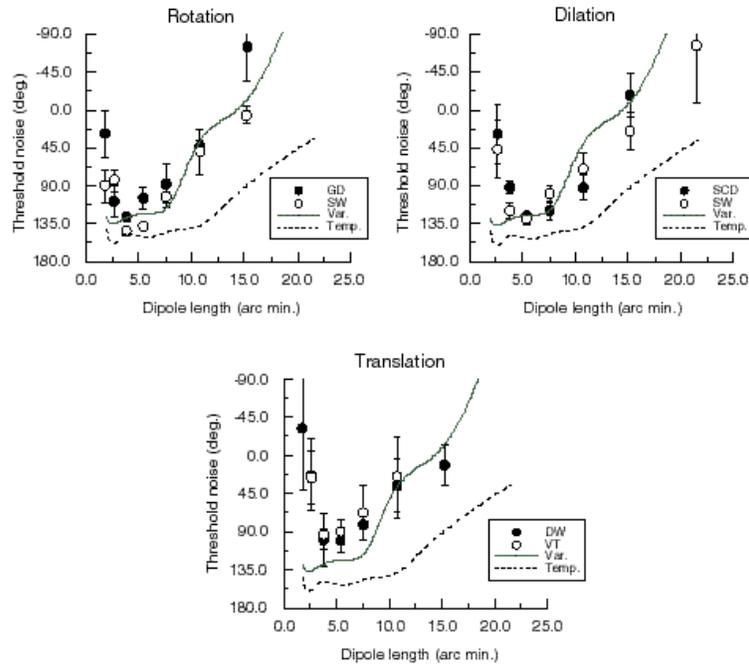


FIGURE 8. Simulation of Experiment 2 using the variance and template models. Note that the performance of both models deteriorates as a function of increasing, but not decreasing, dipole length.

350 The standard deviations of each psychometric function from the simulation are
superimposed on human data in Figure 8. Firstly notice that the template model is very noise-
resistant. Discrimination at separations of around 2.5 to 10 arc min. was achieved when up to
 $\pm 75^\circ$ of variation was present. The best performance of the variance model was discrimination
with approximately $\pm 65^\circ$ of variation, at separations of 2.5 to 6 arc min. The second point of
355 interest is that the models show similar deterioration in performance to human subjects as
dipole separations exceed around 8.0 arc min. No fitting is applied to the predictions, but the
agreement between the variance model, in particular, and human data is good. The final
observation is that neither of the models show any deterioration in performance at very short
dipole lengths. In other words, human observers are unable to make use of the orientation
360 information that is present in patterns containing very short dipoles. This suggests that subjects'
performance is not due to pixelation effects, since both human subjects and models were given
the same stimuli. Approximating a dipole with 1.5 cycles of a square wave, the mechanisms
sensitive to a pattern with a dot separation of 4.0 arc min. are operating at about 23 cycles per
degree. Such mechanisms are known to have much lower contrast sensitivity than mechanisms
365 operating within the middle spatial frequency bands (e.g. using 6 Hz temporally modulated
gratings, contrast sensitivity at 1 c.p.d. is ten times higher than for 20 c.p.d. stimuli; Robson,
1966). Given the limited spatial extent of the stimuli, it is possible that poor performance on
this task is due to a failure of elements to consistently exceed the contrast threshold of the
mechanisms necessary to process them.

370 Notice that slightly poorer performance with translational patterns is reflected by the
poorer fit of the variance model at short dipole lengths. This is interesting because it suggests
that performance differences with transformation type reflect not an advantage for global
organization *per se*, but a *disadvantage* for translational patterns. The present model, based on
local orientation statistics is unable to account for such a disadvantage (albeit a small one).
375 Although performance is better with rotational and dilational patterns, a finding that is
consistent with Wilson *et al.* (1997)'s proposed role for IT, it is interesting to note that subjects'
performance does not *exceed* the performance of a model using local orientation statistics but
merely approaches it. Thus, our data are equally consistent with the presence of some form of

380 low-level interference on the processing of translational patterns as they are with “top-down” facilitation of rotational or dilational Glass pattern structure.

To summarise, this experiment suggests that the representation scheme proposed retains the global orientation structure of texture in the presence of noise, in a manner broadly consistent with human performance. However, the reduced performance of subjects when dealing with orientation structure confined to very high spatial frequency bands (> 23 cycles per degree) is not predicted by the model. This could be corrected by the addition of noise at high spatial frequencies. Both the variance and the template model are unable to account for small observed disadvantage for translational patterns compared to globally organized patterns.

5. EXPERIMENT 3: DETECTION OF ORIENTATION DISRUPTIONS IN GLASS PATTERNS

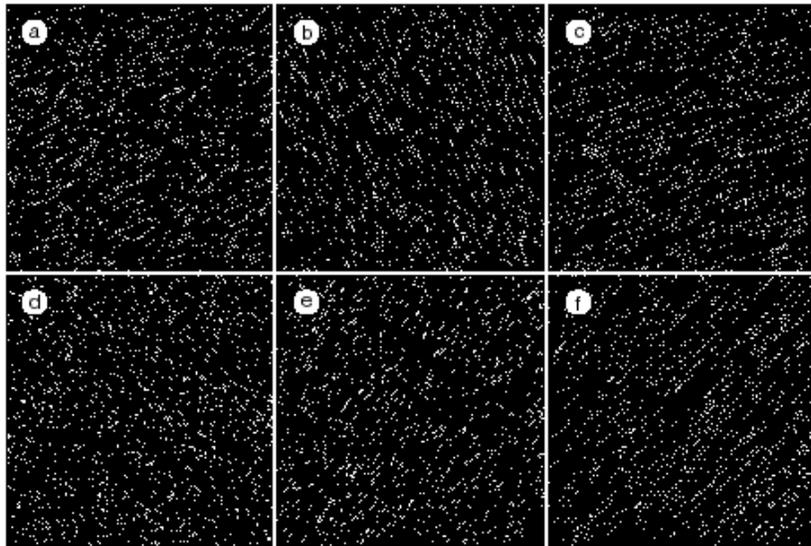
390 In forming a smooth orientation field, some averaging of orientation estimates over space must occur. At first sight it would seem easy to ensure that an orientation field were always smooth by picking very large local regions of integration. This would have two unfortunate side-effects. The first, which was considered above, concerns the smoothing out of structure in globally organized orientation fields. The second is that any local discontinuities in the orientation field would also be smoothed out. Such discontinuities are almost certainly interesting to the visual system as they can signal the occlusion of two surfaces or a surface edge (Hel Or & Zucker, 1989). The questions that this section addresses are: can the proposed model retain interesting discontinuities, and how does it compare to human performance at detecting disruptions to local orientation? In order to establish how accurately discontinuities are retained in human vision, a psychophysical experiment is reported that examines disruptions to local orientation in Glass patterns.

400 Moraglia (1989) showed that detecting a horizontal line element within a circular orientation field is easiest when that element maximally disrupts local orientation. Experiment 3 of Nothdurft (1992) used textures composed of short lines, containing single target lines which differed in orientation to the local orientation of the background field. The required orientation shift, for discrimination of the target from the background, increased as a function of the orientation variation of background elements. Nothdurft (1992) proposes that the detection of anomalous elements is not attributable to the use of feature alignment cues, but instead to the presence of local orientation contrast at the border.

410 The psychophysical procedure used in this section to investigate the detection of orientation discontinuities is an adaptation of a paradigm described in Hel Or & Zucker (1989). They used Glass patterns, as examples of typical densely organized oriented texture, which had small holes inserted in them. These holes were filled with unstructured dots. The threshold size of such embedded patches was determined for discrimination of these textures from ones which did not contain patches. The density of Glass patterns and the “path-length” of each texture element (the number of “overlays” used to generate the pattern) were systematically varied. Hel Or & Zucker (1989) found that the ability to detect anomalous patches was largely independent of the overall density of the pattern. Their claim was that it was primarily *path length* that determined performance. The explanation offered, in terms of curvature constraints on matching, is considered below.

420 There are a number of problems with the psychophysical data presented in Hel Or & Zucker (1989), the main one being that thresholds for 50% correct performance are given. This measure makes the assumption that subjects' behaviour will approach chance *before* the stimulus contains no cue patch, i.e. the psychometric functions will have a measurable “step” around the chance level. In addition, data are only presented for one subject. For these reasons a replication of the original study was performed. The only major change was the use of an adaptive procedure to sample the cue.

425



430 FIGURE 9. Examples of the stimuli used in Experiment 3. The three columns (from left to right) show textures
 435 with a path length of 2, 3 and 6 dots, respectively. Textures in the top row contain patches of unoriented noise
 with radii of 32 pixels, the maximum cue size tested in the experiment. The embedded patch in the bottom row
 has a radius of 16 pixels. Note that the embedded patch becomes progressively easier to detect as the path
 length increases.

435 5.1 Method

The procedure used was similar to the one used in Hel Or & Zucker (1989), except that an adaptive method was used for data collection, making more efficient use of subjects. Unless stated otherwise, stimulus parameters were set to agree with Hel Or & Zucker (1989).

5.1.1 Subjects

440 Two subjects were used in this experiment. One (FJM) was naive to the purpose of the experiment and had normal vision. The other was the author (SCD) who had corrected-to-normal vision.

5.1.2 Apparatus

All stimuli were presented using the same equipment as described in the Methods section above. Subjects were seated in a dimly lit room, 55 cm. away from the monitor.

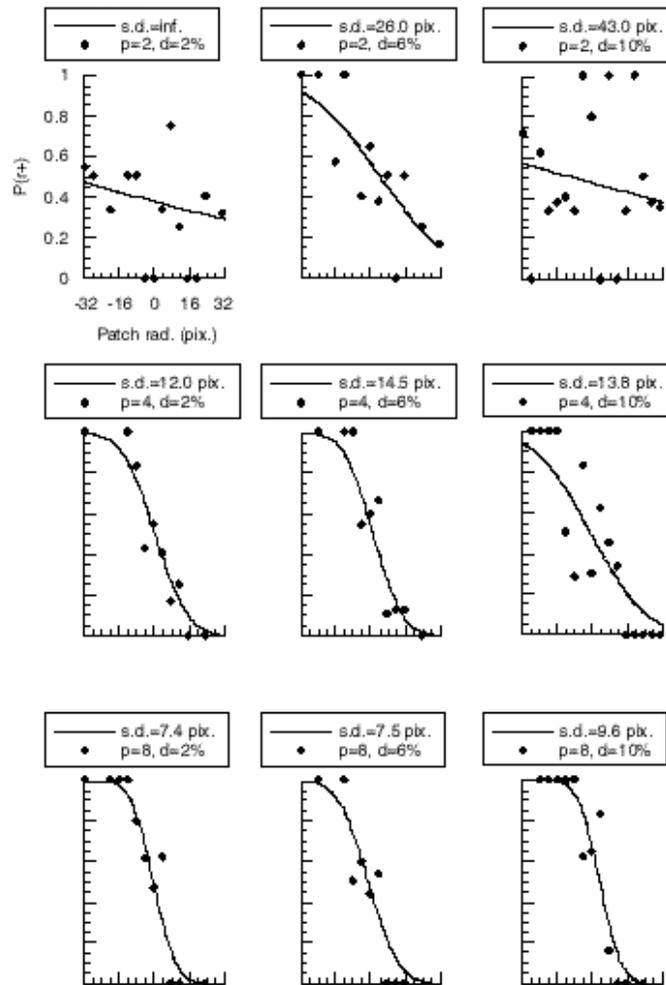
445 5.1.3 Stimuli

The stimuli were translational Glass patterns, created by randomly distributing texture elements, composed of collinear equally spaced dots, within a square field. As the number of dots making up the element (the *path length*) increases, the overall orientation of the pattern becomes progressively clearer (see Figure 9). Dots were composed of single pixels, subtending 2.0 arc min., and appeared white on a black background. The overall orientation of patches was randomised from trial to trial. The separation of dots within each texture element was also randomised within the range 2-4 dot widths (4-8 arc min.). Patterns were presented within 256 by 256 pixel images which were 87 mm square. At the viewing distance of 58 cm, images subtended 8.53 degrees square. Cued textures contained an anomalous embedded field of unoriented dots. This was created by allowing oriented texture elements to fall anywhere within the image except a circular region, with radius r° , creating Glass patterns with small holes. These holes were filled with non-oriented random dot fields, the density of which was identical to the background density of the Glass pattern (precluding the use of a local luminance cue to perform the task). Examples of the stimuli are shown in Figure 9. In order to control for

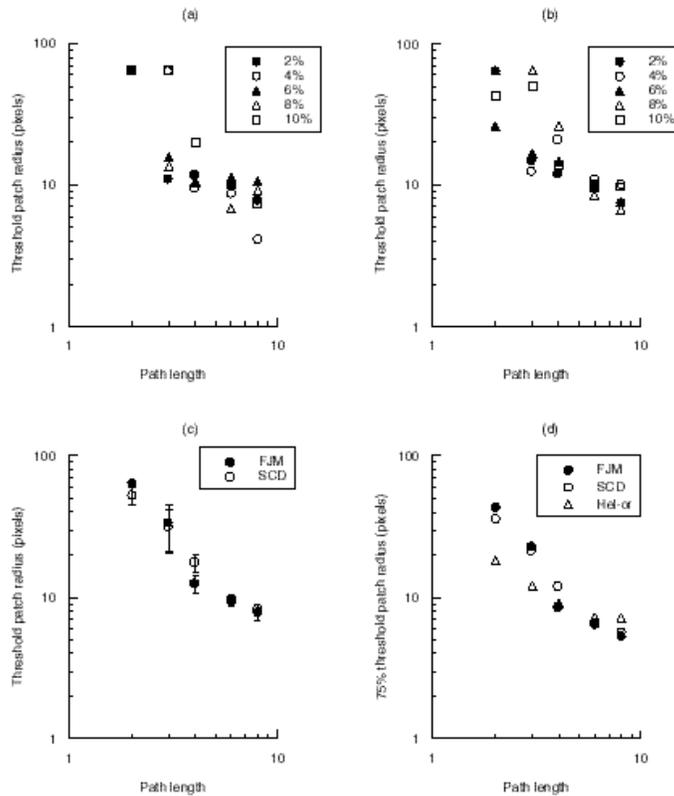
460 the effects of retinal eccentricity the centre of the anomalous patch was randomly located on an
 465 annulus around the centre of the pattern, with radius $(r+2)^\circ$.

5.1.4 Procedure

A two alternative forced choice task was used. This was to decide which of two sequentially
 465 presented Glass patterns contained an anomalous field. The two stimuli were both preceded by
 a fixation marker which was presented in the centre of the screen for 500 ms. An additional 250
 ms of delay followed the offset of the first pattern, making a total ISI of 750 ms. Subjects
 indicated their choice on the computer keyboard. APE (Watt & Andrews, 1981), an adaptive
 method of constant stimuli, was used to determine the threshold patch size for discrimination.
 470 A range of patch radii from 0.0 to 32.0 pixels (64 arc min.), were adaptively sampled in steps of
 4.0 pixels (8 arc min.). Runs of 64 trials were used with no interleaving. Dot densities of 2, 4, 6, 8
 and 10% were used and the path lengths tested were 2, 3, 4, 6 and 8 dots. Conditions were
 blocked separately in all cases.



475 **FIGURE 10.** Typical psychometric functions for the anomalous patch detection task. Axes legends are shown in
 the top left-hand graph. In the legend boxes p refers to the patch length and d the flow density. The sign of the
 480 abscissa, the patch size, indicates whether the cued texture was presented in the first interval (negative) or
 the second interval (positive). The ordinate shows the probability that the response was that the cued texture
 was in the first interval. Note that fits to the data shown, using the standard cumulative Gaussian
 psychometric function, provides a reasonable fit to most of the data and that there is no sign of a plateau in
 performance around the 50% level (as presumed in Hel Or & Zucker, 1989)



485 **FIGURE 11.** Data from Experiment 3. (a-b) Threshold patch radius for detection of an anomalous patch for two subjects (FJM and SCD respectively) as a function of path length. Note that data from all densities shows the same trend. (c) Data from both subjects averaged across densities. Note the steady decrease in required patch size as path length increases. (d) Comparison of subjects detection of anomalous patches from this and Hel Or & Zucker (1989). Data are re-expressed as 50% thresholds and averaged over all density conditions. The agreement is fairly good for thresholds greater than 8.0 pixels. The lower thresholds in this study are probably due to the fact that adaptive sampling of patch sizes below 8.0 pixels was used.

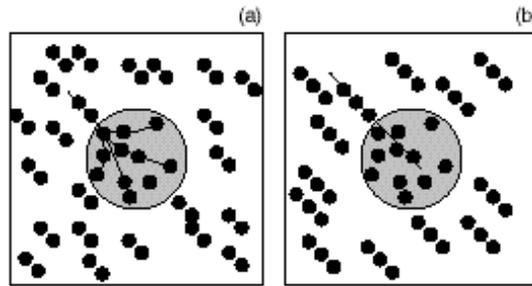
490

5.2 Results

495 The major difference between this and the Hel Or & Zucker (1989) study is the use of an adaptive procedure, rather than a method of constant stimuli, and subsequent measurement of thresholds at the 83% and not the 50% level. A sample of the psychometric functions from one subject is shown in Figure 10 along with fits from the standard cumulative Gaussian

500 psychometric function. Thresholds are presented for both subjects in Figure 11 as a function of path length. The patch size required for discrimination falls steadily as a function of path length. Density does not seem to affect the shape of this function and for this reason data are also presented averaged across density in Figure 11c. In order to compare data from this experiment with those from the original study, thresholds for the 50% level were also calculated. Results from two subjects are shown in Figure 11d, along with data from Hel Or & Zucker (1989). The first set of plots shows the minimum patch size required to perform at chance (50%) level as a function of path length and texture density. The agreement between the original study and this experiment seems reasonable. Both sets of data show that the minimum patch size for performing the task depends primarily on path length and not density.

505



510

FIGURE 12. Hel Or & Zucker (1989) curvature hypothesis, adapted from their Figure 4. (a) Shows a flow pattern with path length two. There are many potential matches between elements near the patch and elements in the patch because paths are not constrained by curvature. (b) A flow with path length three. Curvature limits potential matches and increases the probability of detecting the patch.

515

5.3 Hel Or & Zucker (1989) interpretation

In order to propagate and smooth the orientation field given by the output of a set of receptive fields, Hel Or and Zucker propose that an interactive, cooperative procedure is used. This refinement occurs through inter-columnar interaction in V1 (Parent & Zucker, 1989) and operates by modifying the firing rate of cells within a neighbourhood according to their mutual consistency with respect to orientation. This amounts to propagating curves within the image. However, at the border of a texture patch orientation information is ambiguous and, Hel Or and Zucker claim, curvature information is necessary to stop the smoothing out of the discontinuity.

520

525

Figure 12 illustrates this explanation. The left hand part of the figure shows a Glass pattern with a path length of two. At the patch border a wide range of possible matches are possible, because matches are unconstrained with respect to curvature. There will be enough random matches that match the background orientation to ensure that smooth patches will be smoothed out. For the case of a Glass pattern with path length three, illustrated in Figure 12b, matches at the border are constrained to have zero curvature and so many fewer matches are possible. Consequently, patches are less likely to be smoothed out.

530

There are a number of problems with this explanation. Firstly the result that it is hard to detect anomalous patches in textures with a path length of two does not necessarily mean that recourse need be made to an explanation in terms of curvature. It is quite possible that an extra dot in the path merely increases the differential output of a filter at a particular scale/orientation and it is this increase in energy, relative to the output over the anomalous patch, that improves performance. Since the curvature model has not been implemented its adequacy has not been tested quantitatively. An alternative explanation using orientation reliability is given in the next section.

535

540

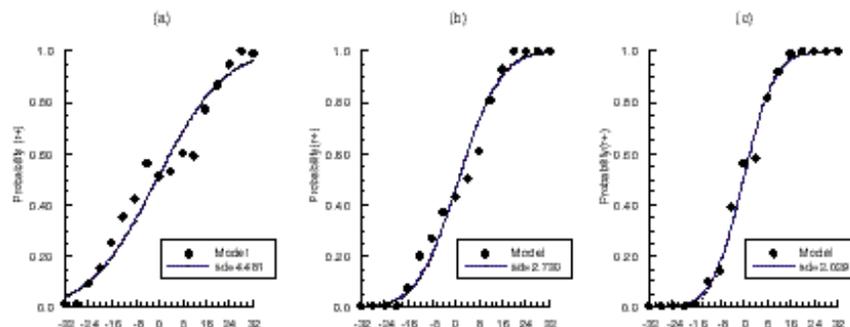


FIGURE 13. Typical psychometric functions from the simulation. This figure shows the performance of the discriminator using mean local orientation variance for textures with densities of 6% and path lengths of (a) 2, (b) 4 and (c) 8 dots.

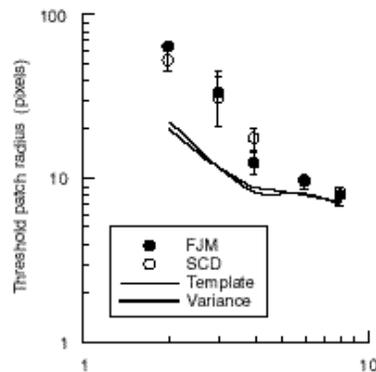
545

5.4 An alternative interpretation

550 In this experiment, anomalous patches were marked out from the background by the fact that they were unoriented. We would expect such a patch to be signaled by a region of high local orientation variance. It is possible this source of local orientation variance will elevate the estimate of multi-local variance for the whole pattern sufficiently to permit detection of the patch. Alternatively, the unstructured patch will produce a deviation from expected local
555 orientation which will be measurable by the template estimate. To compare these strategies we used identical estimates of multi-local variance and template-match to those described above.

5.5 Details of simulation

A simplified version of the model was again used, due to processing time restrictions. The
560 average optimal spatial scale was calculated over 32 Glass patterns at the beginning of a run and subsequent convolutions were only performed *at this scale*². The same five dot densities (2%, 4%, 6%, 8% and 10%) and path lengths (2, 3, 4, 6 and 8 dots) were used as in the psychophysical experiments, giving a total of 25 psychometric functions to generate. For each run, 32 stimuli were generated at each of nine different patch radii from 0.0 to 32.0 pixels.
565 Having collected measurements of structure from the model at each of the cue levels, a psychometric function was calculated by sampling at each stimulus level and calculating the probability of discrimination given the models' output from a cued pattern compared to the output given a reference pattern.



570 **FIGURE 14.** Results from the simulations of the task from Experiment 3. Predictions from the variance and template models are consistent with human subjects.

5.6 Simulation results

575 Figure 14 shows predictions from the variance and template model for this task. The fit to human data is remarkably good given that each point represents the standard deviation of an entire psychometric function and that no fitting has been applied.

The success of the variance model strongly suggests that this task may not require any kind of boundary detection as supposed. On inspection, boundaries can be detected within the image, but the overall smoothness of the texture proves to be an adequate indicator of the presence of a patch. In order to preclude the use of simple global statistics in performing segmentation tasks, some authors (e.g. Landy & Bergen, 1991; Nothdurft, 1985) use patch *shape* discrimination tasks, and a thorough comparison of these models for texture segmentation would certainly have to use data from studies such as these.

585 These simulation results are in broad agreement with those of Nothdurft (1992), since the variance model would also predict that background orientation variation will critically affect

² Even using this simplification, the simulation described required approximately 92000 convolutions which took 15 days on a Hewlett Packard series 700 Apollo work station.

detection of anomalous elements. Nothdurft suggests that texture disruption is not detected by breaks in the co-aligned texture elements but in local orientation contrast at the anomalous region. Using the pre-processing stage proposed here, breaks in local feature alignment would be coded by a failure to coalesce into single elements at coarser spatial scales. Pilot trials demonstrated that this was not a reliable cue to the presence of an anomalous patch. Instead the proposed (variance) model codes the sum of local orientation variance. This measure covaries with the degree of local orientation contrast.

6. GENERAL DISCUSSION

This paper has reported a comparison of an adaptive filtering model combined with a measure of mean local orientation variance, to one using a measure of the degree of local orientation matching to a template. The intention was to test the spatial aspects of the model by considering how globally organized patterns are coded. It has been shown that when these measures are treated as discriminators they approach and exceed human levels performance on three different psychophysical tasks, involving the perception of structure in Glass patterns. As is often the case however, the model's failures prove to be more illuminating than its successes.

Data from Maloney *et al.* (1987) (Experiment 1) demonstrated that while a measure of multi-local orientation variance could predict the discriminability of Glass patterns in the presence of extra noise dots there had to be sufficient texels present for the measure to be of any use. At low texel densities it was suggested that subjects make use of more detailed knowledge about the underlying transform. This suggests that local feature density may prove to be a critical determinant of the type of texture processing that an image undergoes.

Experiment 2 demonstrated that rotational textures are less vulnerable to the intrusion of orientation noise than translational textures. This is consistent with a recent report by Wilson *et al.* (1997) who attribute this phenomenon to the operation of concentric structure detectors in IT. The simulation data reported suggest that what is happening is that higher-level structure is providing access to exactly the levels of performance that would be predicted from the use of local orientation statistics. It seems that what we are observing is a *translational disadvantage* rather than a *concentric advantage*; how this can be reconciled with either global or local models has yet to be determined.

The Hel Or & Zucker (1989) task replicated and modeled in Experiment 3, is closely related to a variety of texture segmentation tasks. The success of the orientation variance model at accounting for human performance suggests that this task does not require the explicit coding of the *position* of texture boundaries at all, and that instead local orientation statistics suffice. I consider that this is illustrative of a general problem with texture segmentation tasks - the information that subjects can use is often insufficiently constrained. In fact, the location of texture boundaries may ultimately prove to be a visual strategy that is little more than a byproduct of a texture processing scheme intended primarily to guide which filters are used for what.

7. REFERENCES

- Bergen, J. (1991). Theories of visual texture perception. In D. Regan (Ed) *Vision and visual dysfunction.*, vol. 10: Spatial Vision (pp. 114-134). London: MacMillan.
- 630 Dakin, S. C. (1997a). The detection of structure in Glass patterns: Psychophysics and computational models. *Vision Research*, 37, 2227-2259.
- Dakin, S. C. (1997b). Glass patterns: Some contrast effects re-evaluated. *Perception*, 26, 253-268.
- Dakin, S. C. and Watt, R. J. (1997). The computation of orientation statistics from visual texture. *Vision Research*, 37, 3181-3192.
- 635 Dodwell, P. and Caelli, T. (1985). Recognition of vector patterns under transformation: Local and global determinants. *Quarterly Journal of Experimental Psychology*, 37A, 1-23.
- Freeman, W. H. and Adelson, E. H. (1991). The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13, 891-906.
- Glass, L. (1969). Moiré effects from random dots. *Nature*, 243, 578-580.
- 640 Hallet, P. (1992). Segregation of mesh-derived textures evaluated by resistance to added disorder. *Vision Research*, 32, 1899-1911.
- Hel Or, Y. and Zucker, S. (1989). Texture fields and texture flows: Sensitivity to differences. *Spatial Vision*, 4, 131-139.
- Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature*, 290, 91-97.
- 645 Landy, M. S. and Bergen, J. R. (1991). Texture segregation and orientation gradient. *Vision Research*, 31, 679-91.
- Maloney, R., Mitchison, G. and Barlow, H. (1987). Limit to the detection of Glass patterns in the presence of noise. *Journal of the Optical Society of America*, A4, 2336-2341.
- 650 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, California.: Freeman.
- Moraglia, G. (1989). Display organisation and the detection of horizontal line segments. *Perception and Psychophysics*, 45, 265-272.
- 655 Nothdurft, H. (1992). Feature analysis and the role of similarity in preattentive vision. *Perception and Psychophysics*, 52, 355-375.
- Nothdurft, H. C. (1985). Sensitivity for structure gradient in texture discrimination tasks. *Vision Research*, 25, 1957-68.
- Parent, P. and Zucker, S. W. (1989). Trace inference, curvature consistency and curve-detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 823-839.
- 660 Phillips, G. and 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.
- 665 Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1992). *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge, UK.: Cambridge University Press.
- Robson, J. (1966). Spatial and temporal contrast sensitivity functions of the visual system. *Journal of the Optical Society of America*, 56, 1141-1142.
- Swets, J. (1964). *Signal detection and recognition by human observers*. New York, NY: John Wiley and Sons.
- 670 Watt, R. J. and Andrews, D. (1981). APE: Adaptive probit estimation of psychometric functions. *Current Psychological Review*, 1, 205-214.
- Westheimer, G. (1981). Visual hyperacuity. *Progress in Sensory Physiology*, 1, 1-20.
- Wilson, H. R., Wilkinson, F. and Asaad, W. (1997). Concentric orientation summation in human form vision. *Vision Research*, 37, 2325-2330.
- 675 Zucker, S. (1982). Early orientation selection and grouping: Evidence for type I and type II processes. Montreal, Quebec.: McGill university.

680 **8. Acknowledgments** - I am grateful to Roger Watt and Ian Paterson for their considerable contribution to this work including discussion and the provision of software. Thanks also to Fred Kingdom for commenting on an earlier draft of this work

9. This work was supported by the SERC (grant GR/H53181) while the author was at the University of Stirling.

10. APPENDIX: A MODEL FOR ESTIMATING MULTI-LOCAL ORIENTATION STATISTICS

685 This appendix describes a model for extracting and representing the orientation of features derived from texture. Table 2 summarises the models parameters and the values used in the simulations reported.

| Parameter | Symbol | Value |
|---|--------|--|
| No. of scales | | 7 |
| Filter size (spatial frequency) | | 1-8 pixels in 0.5 octave steps (40.7-5.0 cycles per image) |
| No. of filter orientations | | 12 |
| Filter orientations | | 0°-135° in 15° steps |
| Gaussian smoothing parameter | | 2 |
| Threshold to demarcate blobs | | 1 grey-level standard deviation |
| Neighbourhood sampling rate (Sampling grid sizes) | | 32/ (32X32 to 4X4) |
| Neighbourhood size | R | pixels |

TABLE 2. Summary of model parameters.

690 **10.1 Estimation of feature orientation**

Pre-processing consists of convolution with a bank of 2D, oriented, Difference-of-Gaussian (DoG) filters. This filter is composed of a DoG in the x-direction multiplied by a Gaussian in the y-direction:

695
$$W(x, y) = \exp\left(-\frac{x^2}{2\sigma^2}\right) - \frac{1}{2.23} \exp\left(-\frac{x^2}{2(2.23\sigma)^2}\right) \exp\left(-\frac{y^2}{2(3\sigma)^2}\right)$$
 (5)

where σ refers to the standard deviation of the positive Gaussian function and (x_i, y_i) are co-ordinates rotated by angle θ :

700
$$x = x \cos \theta + y \sin \theta$$
 (6)

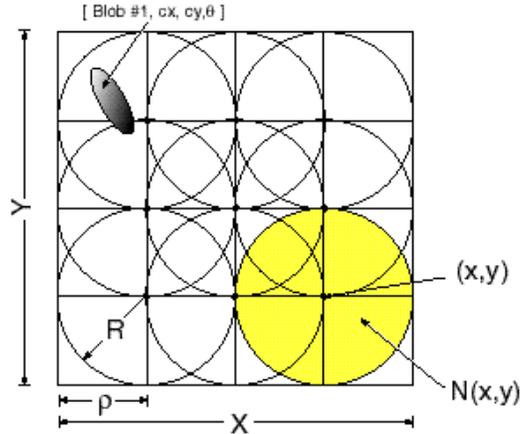
$$y = y \cos \theta - x \sin \theta$$
 (7)

Parameters of the DoG are based on those derived by Wilson and co-workers (e.g. Phillips & Wilson, 1983). The model uses filter sizes of $2^{(k+1)/2}$ pixels, $k = 1, 2, \dots, 7$, and orientations of $\theta = 0-165^\circ$ in steps of 15° .

705 Next, the squared filter output is computed across all orientations for each spatial scale, and a small amount of Gaussian blur is applied to each of these “local energy” representations. A new image is constructed by comparing, pixel by pixel, local energy across orientation and assigning a value to the pixel equal to the original output of the filter who’s energy is maximal
710 (see Figure 2b for examples). Thus the system selects the most appropriate filter based on local

image structure. Although implemented in a quite different manner it is in this regard similar to a “steerable” filter system (Freeman & Adelson, 1991).

715 After the initial filtering stage a symbolic, “primal sketch” description of thresholded blobs is constructed. Filter outputs are doubly half-wave rectified, and positive and negative zero-bounded regions (or “blobs”) characterized using the image description scheme of Watt (1991). This produces “sentences” describing blobs in terms of their principal axis, centroid position, area, etc. The orientations of these blobs serve as input to the next stage of processing.



720

FIGURE 15. Image partitioning scheme used for calculation of local orientation statistics.

10.2 Partitioning the image

Given a set of orientation estimates at scale ρ , denote the image lattice (the set of all pixel positions) as:

725

$$I = \{(i, j) : 1 \leq i, j \leq N\} \quad (8)$$

The resolution of analysis determines the sampling rate of a sublattice of points $I' \subset I$:

$$I' = \left\{ \left((i-0.5) \frac{X-1}{N-1}, (j-0.5) \frac{Y-1}{N-1} \right) : 1 \leq i \leq X, 1 \leq j \leq Y \right\} \quad (9)$$

730

Now let $N(x, y)$ define a circular neighbourhood, radius R , around each point on the sublattice, I' , (see Figure 15):

$$N(x, y) = \left\{ (i, j) : (x-i)^2 + (y-j)^2 \leq R^2 \right\} \quad (10)$$

735

where $(x, y) \in I'$. This defines a set of circular neighbourhoods distributed throughout the image on a grid. The two parameters used in this section are ρ the sampling rate of the sublattice, and R the radius of the local region of integration. The values of these parameters are constrained firstly by the size of the filters measuring local orientation estimates and secondly by the requirement of the model to be *scale-invariant*.

740

The first constraint is that it is meaningless to use a sampling rate or a neighbourhood size which is too small in comparison to the filter size ρ . The maximum sampling rate was set to be $S_{1.0} = 8.0$ pixels (i.e. a 32×32 sublattice) at the finest spatial scale ($\rho = 1.0$ pixel) and $R = \rho$ as neighbourhood size.

745

The second constraint, scale invariance, has been demonstrated psychophysically by Nothdurft (1985). Subjects' discrimination of texture pairs is impaired if the spacing of elements is increased but micro-pattern size remains constant. If both are increased *proportionally*

segregation is not affected. In effect this requires that the sampling rate be reduced in proportion to the spatial scale of analysis:

$$750 \quad \frac{1}{\dots} \quad (11)$$

And I used:

$$= \frac{32}{\dots} \quad (12)$$

755 Features in the image are partitioned using their “blob” characteristics from the sentence based description produced by earlier processing (see Section 8.1). Blob position is given by:

$$C = \{ [cx_{,k}, cy_{,k}], 1 \leq k \leq B \} \quad (13)$$

760

where $(cx_{,k}, cy_{,k})$ is the centroid of the kth blob, and B is the total number of blobs at scale d . The set of all blobs falling in a neighbourhood around (x, y) is:

$$765 \quad C(x, y) = N(x, y) \quad C \quad (14)$$

This defines the sets of blobs that are used to calculate local texture statistics. The following two sections describe how the mean and variance of blob orientations are calculated.

10.3 Multi-local orientation statistics

770 10.3.1 Local mean orientation

The orientation of a blob with centroid (x, y) is denoted $\theta_{(i,j)}$. The basic expression for estimating the mean orientation of the set of blobs falling in $C(x, y)$:

$$\bar{\theta}(x, y) = \frac{1}{2} \tan^{-1} \frac{\sum_{(i,j) \in C(x,y)} \sin 2\theta_{(i,j)}}{\sum_{(i,j) \in C(x,y)} \cos 2\theta_{(i,j)}} \quad (15)$$

775 This expression is ambiguous; the result is either the mean, or the mean offset by 90° . To resolve this ambiguity we assess the second derivative of the underlying error function:

$$\frac{d^2 E}{d^2} = -2 \sum_{(i,j) \in C(x,y)} \cos 2(\theta_{(i,j)} - \bar{\theta}(x, y)) \quad (16)$$

780

To give the final expression:

$$\bar{\theta}(x, y) = \begin{cases} \bar{\theta}(x, y) & \text{if } \frac{d^2 E}{d^2} < 0 \\ \bar{\theta}(x, y) + 90^\circ & \text{otherwise} \end{cases} \quad (17)$$

With $\bar{\theta}(x, y)$ being coded between 0 and 180° (For derivation see Dakin, 1997)

785 10.3.2 Local orientation variance

A useful measure of orientation smoothness is the mean of a set of estimates of local orientation variance:

$$= \frac{1}{XY} \sum_{(i,j) \in C} \sigma^2(i,j) \quad (18)$$

790 Where X and Y are the image dimensions, σ^2 is the sampling rate, and $\sigma^2(x,y)$ is a measure of the orientation variance in subregion $C(x,y)$, at scale σ :

$$\sigma^2(x,y) = \frac{1}{|C(x,y)|} \sum_{(i,j) \in C(x,y)} \sin^2(\theta(i,j) - \bar{\theta}(x,y)) \quad (19)$$

where $|C(x,y)|$ is the number of blobs in each subregion.

795 While this scheme produces an estimate of the smoothness of the orientation field it gives no indication of the global organisation of the underlying field, i.e. it cannot be used to distinguish a rotational field from a translational one. Presumably this is a natural task for subjects, and for that reason we consider a second measurement of image structure: a template match.

800 10.3.3 Local template matching

If one has prior knowledge of the possible underlying orientation structure of an image a simplest method to assess the "degree of structure" is to calculate the deviation of local mean orientations $\bar{\theta}(x,y)$ from the expected orientation at that position, $T(x,y)$:

$$D = \frac{1}{XY} \sum_{(i,j) \in I} (\bar{\theta}(i,j) - T(i,j))^2 \quad (20)$$

805 and to sum over all regions in the image:

For a rotation (centre (x_c, y_c)), for example, the expected value is defined as:

$$T_{rot}(x,y) = \tan^{-1} \frac{y_c - y}{x_c - x} + 90^\circ \quad (21)$$

810 for a dilations as:

$$T_{dil}(x,y) = \tan^{-1} \frac{y_c - y}{x_c - x} \quad (22)$$

and for a translation, angle A as:

$$T_{trans}(x,y) = A \quad (23)$$

815 The obvious disadvantage to this method is that the underlying transformation must be known *a priori* in order to calculate $T(x,y)$. However, given that for many psychophysical tasks different types of patterns are not interleaved in a single run, it may be assumed that subjects often have this information.

820