A Linear Systems Approach to the Detection of Both Abrupt and Smooth Spatial Variations in Orientation-defined Textures

FREDERICK A. A. KINGDOM,* DAVID R. T. KEEBLE*

Two distinct paradigms have characterized most previous studies of texture perception: one has dealt with texture *segregation*, the other with the processing of texture *gradients*. Typically, studies of texture segregation have used stimuli with abrupt textural variations, whereas studies of texture gradient processing have used stimuli with smooth textural variations. In this study we have asked whether the mechanisms which process abrupt and smooth textural variations are the same, by considering whether a simple linear model can account for the detection of orientation modulation in micropattern-based textures with three types of modulation: sine-wave (SN), square-wave (SQ) and missing fundamental (MF). The MF waveform was constructed by removing the fundamental harmonic from a square-wave. We found a clear overall ordering of sensitivity: SQ > SN > MF. We found that sensitivity to the SQ and MF stimuli could be predicted very well from the SN data if one assumed that the r.m.s. output of a single linear channel underlay the detection of the orientation modulation. This suggests that the detection of both abrupt and smooth changes in orientation-defined textures is subserved by a common mechanism which mimics the operation of a single linear channel.

Texture   Orientation   Linear systems analysis

INTRODUCTION

The study of texture perception is important because textural variations across space provide important information about both the location of objects and the shapes of their surfaces. Past research into texture perception has, broadly speaking, fallen into two categories. One category has examined how we *segregate* textural regions which differ along one or more textural dimension, and this is perhaps best exemplified by the early studies of Beck (1966) and Julesz (1975), as well as a large number of more recent studies (e.g. Beck, 1982; Gurnsey & Browse, 1988; Vorhees & Poggio, 1988; Rubenstein & Sagi, 1990, 1993; Nothdurft, 1991; Malik & Perona, 1990; Landy & Bergen, 1991; Graham, Beck & Sutter, 1992; Lamme, Van Dijk & Spekreijse, 1992; Gorea & Papathomas, 1993; and see review by Bergen, 1991). A second category has examined how we process texture *gradients*, and this is principally associated with the early seminal work of Gibson (1950), as well as numerous more recent studies (e.g. Cutting & Millard, 1984; Todd & Akerstrom, 1987; Stevens & Brookes, 1988; Buckley & Frisby, 1993; Blake, Bulthoff & Sheinberg, 1991).

The computational significance of texture gradients for shape-from-texture processing has also attracted much interest (Witkin, 1981; Davis, Janos & Dunn, 1983; Kanato & Chou, 1989; Blake & Marinos, 1990; Linderburg & Garding, 1993). It is reasonable to ask whether a common mechanism underlies the detection of the abrupt changes in stimuli employed in studies of texture segregation and the smoothly varying changes which characterize stimuli in studies of texture gradient processing. A recent study by Kingdom, Keeble and Moulden (1995) began to address this issue with respect to orientation-defined texture perception. Kingdom *et al.* employed a stimulus consisting of a dense array of Gabor micropatterns whose positions were randomized but whose orientations were varied systematically across the display according to a specified periodic waveform. The principal descriptor of performance was the orientation modulation function, which described the sensitivity to the orientation modulation as a function of the spatial frequency of the orientation modulation. One experiment in their study compared orientation modulation functions for sine-wave and square-wave orientation modulations. The sine-wave texture exemplified a stimulus with smooth texture variations whereas the square-wave stimulus exemplified one whose textural variations were abrupt. Kingdom *et al.* found that the orientation modulation functions for the sine- and

*McGill Vision Research Unit, Department of Ophthalmology, Room H4-14, 687 Pine Avenue West, Montreal, Quebec, Canada H3A 1A1 [*Email* fred@jiffy.vision.mcgill.ca].
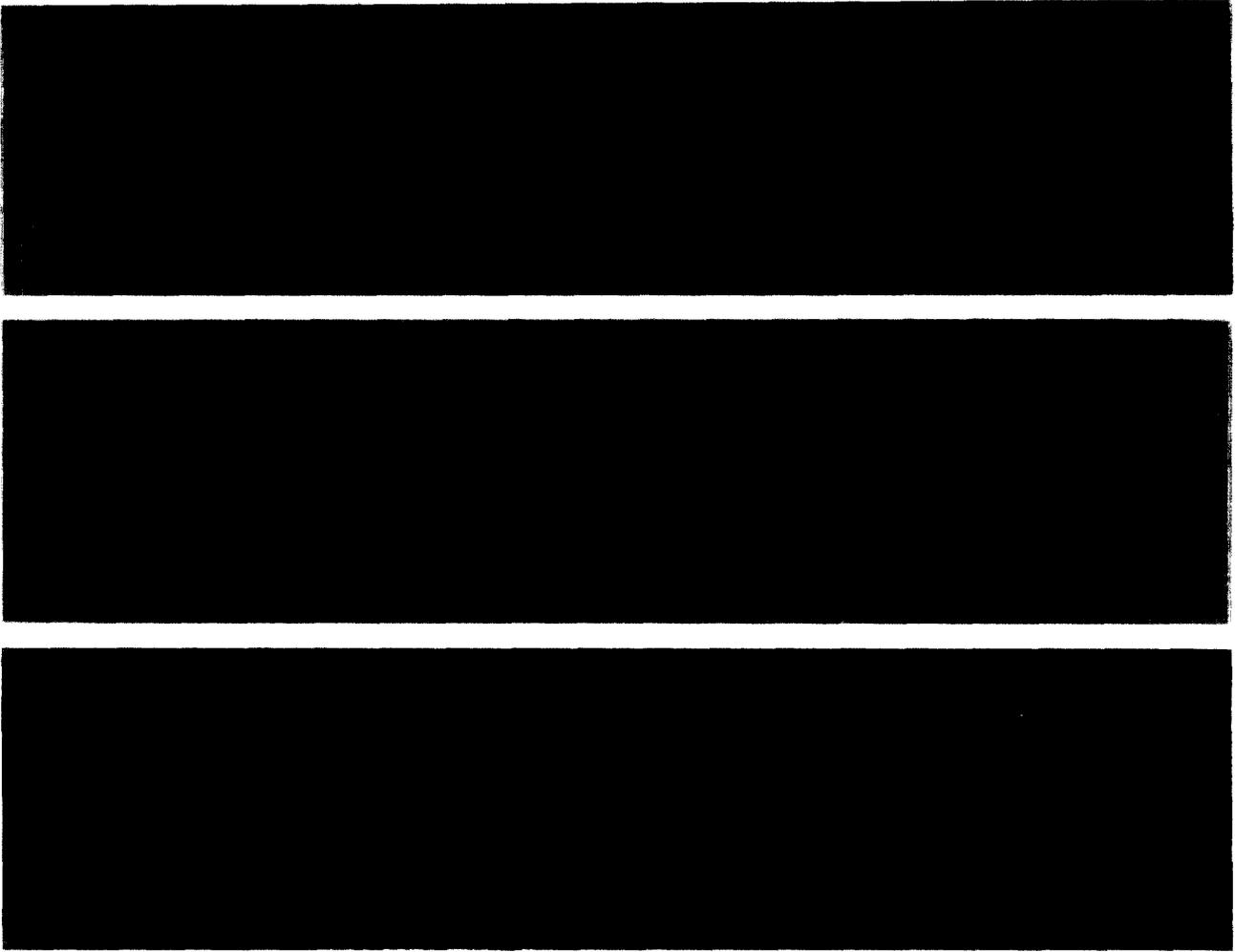
FIGURE 1. Examples of the stimuli, with amplitudes of orientation modulation greater than that used in the experiments for illustrative purposes.

square-wave stimuli were quite similar in shape, with the square-wave to sine-wave sensitivity ratio being about 1.34 on average. They concluded that the similarity in shape of the sine- and square-wave orientation modulation functions was consistent with a common underlying mechanism.

In this study we extend the initial findings of Kingdom *et al.* with sine- and square-wave stimuli by comparing the detectability of three types of waveform: sine-wave (SN), square-wave (SQ) and missing fundamental (MF), at a number of spatial frequencies of orientation modulation, and using a large number of subjects (six). Examples of the actual stimuli are shown in Fig. 1 and a schematic representation of the underlying waveforms is shown in Fig. 2. The missing fundamental is generated by subtracting the fundamental harmonic component of a square-wave from the square-wave itself (Campbell, Howell & Robson, 1971). Comparison of the SN, SQ and MF stimuli allows us to consider whether the detection of different forms of orientation modulation is governed by a common linear model. Campbell and Robson (1968) compared the detectability of just SN and SQ *luminance*-modulated gratings at various spatial frequencies and the result they obtained was the first decisive piece of psychophysical evidence in support of
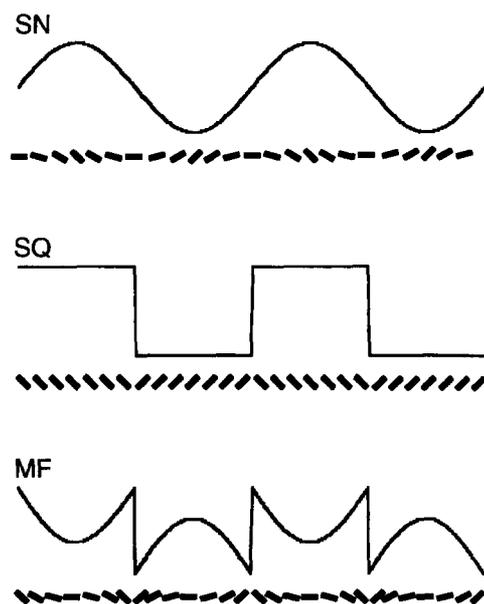
FIGURE 2. Schematic representation of the stimuli used in the experiments. Each waveform describes how the orientations of the component Gabor micropatterns change across the stimulus. Below each waveform are shown diagrammatic versions of the micropatterns with their approximate orientations for a given amplitude of orientation modulation, though without the orientation "noise" added to the micropatterns in the actual stimuli used in the experiments.

the view that luminance vision at threshold is subserved by a set of independent linear channels tuned to different spatial frequency ranges. This hypothesis received additional support in later studies by Campbell et al. (1971) and Campbell, Howell and Johnstone (1978), which compared the detectability of, and measured the discriminability between, SN, SQ and MF, luminance-defined stimuli. In all these studies, the critical test of whether a linear model underlay detection was whether the data for the SN stimuli could predict without any free parameters the detectability of the SQ and MF stimuli. We have applied the same logic to the detection of orientation modulation in our textured stimuli.

In summary, this study aims to determine whether a common linear mechanism exists for detecting both abrupt and smooth changes in orientation-defined textures, by considering whether the measured sensitivity to SN-modulated textures can predict the measured sensitivity to SQ- and MF-modulated textures. A brief report of the results of this study has been given elsewhere (Kingdom & Keeble, 1994).

METHOD

Subjects

Six subjects were employed in the experiments. FK and DK were highly experienced psychophysical observers, while MV, AW, LM and SA were paid volunteers who were naive as to the purpose of the experiments. All had normal or corrected-to-normal vision.

Stimuli

Generation. The stimuli were generated by a Macintosh IIfx computer and displayed on a SuperMac Trinitron monitor. The display was monochrome and was gamma corrected by suitable selection of intensity levels from an 8-bit (256 grey levels) look-up-table following calibration using a UDT (United Detector) photometer.

Gabor micropatterns. The function for each Gabor micropattern was defined as a one-dimensional cosine function multiplied by a two-dimensional isotropic Gaussian envelope

$$L(x,y,\theta) = M + A\cos[2\pi f(x\cos\theta - y\sin\theta)]$$
$$\times \exp[-(x^2+y^2)/2\sigma^2)]. \quad (1)$$

In equation (1) M is mean luminance, A luminance amplitude, f spatial frequency, σ the space constant of the Gaussian envelope and θ micropattern orientation. The centre spatial frequency, f, of the Gabor micropatterns was 4.8 c/deg when viewed at the standard distance of 63 cm. σ was set equal to 0.125 deg, giving the micropatterns a spatial frequency bandwidth at half-height of 0.94 octaves. The Gabor micropatterns all had a contrast of 39%, where contrast was defined as the peak amplitude, A, divided by the mean, M. Both M and the background were set to 34 cd m^{-2}.

SN, SQ and MF orientation-modulated textures. The SN and SQ waveforms were defined conventionally.

The MF waveform was obtained by subtracting the fundamental harmonic component from a square-wave. The fundamental harmonic of a square-wave is a sine-wave with an amplitude of $4/\pi$, or 1.273, times that of the square-wave itself. The three types of waveform are shown schematically in Fig. 2. The corresponding orientation of the micropatterns is shown below each point on the waveform.

Examples of the stimuli employed are shown in Fig. 1, although for illustrative purposes they are shown with higher amplitudes of orientation modulation than used in the actual experiments. Each stimulus was an array 30 × 7 deg containing 1000 micropatterns. The positions of the individual micropatterns within the display window were randomized. The orientation of the micropatterns was constrained in that the nominal mean orientation of the micropatterns varied along the horizontal axis of the display according to the specified waveform (SN, SQ or MF), with that modulation being about the horizontal. The amplitude of orientation modulation was determined by how much the orientation of the micropatterns changed throughout one complete cycle of orientation modulation. For example, an amplitude of orientation modulation of 10 deg implied that the micropatterns changed by 20 deg throughout one complete cycle of orientation modulation. For each waveform, amplitude was defined as half the peak-to-trough difference in orientation. All micropatterns at a given horizontal location were given an orientation drawn randomly from a Gaussian distribution of orientations with a specified mean (determined by the point on the waveform) and SD. The SD of the Gaussian distribution of orientations was 10 deg, and this represented the amount of orientation "noise" in the stimulus. Where Gabor patches overlapped, they were combined additively.

Procedure

A two-interval forced-choice (2IFC) paradigm was used in all experiments to measure the threshold amplitude of orientation modulation. On each trial two displays were presented, each for 107 msec, with a 2 sec inter-stimulus-interval. There was one exception, subject LM, for whom the stimuli were displayed for 200 msec. LM complained during practice trials that the 107 msec exposure duration felt uncomfortable, and because her performance was also particularly poor at this duration, it was decided to increase her exposure duration to 200 msec. The task for the subjects on each trial was to decide which interval contained the stimulus with the orientation modulation. The only difference between the two stimuli presented on a given trial was in their amplitude of orientation modulation, which in the comparison stimulus was zero. The method of constant stimuli was used with five amplitudes of orientation modulation, the magnitudes of which were determined for each spatial frequency of orientation modulation by pilot studies. During a given session, which consisted of 200 trials, only one type of waveform was presented, but the five spatial frequencies of orientation modulation, at

each of their five amplitudes, were presented in random order. There were seven sessions conducted for each waveform, resulting in a total of 280 trials for each psychometric function, and thus for each data point. The five spatial frequencies were 0.033, 0.067, 0.133, 0.267 and 0.533 c/deg, and for the fixed display size these resulted in 1, 2, 4, 8 and 16 cycles of orientation modulation respectively. The phase of orientation modulation was also randomized on each trial. Sessions with different waveforms were presented in random order. Feedback on each trial was given in the form of a tone for an incorrect decision.

RESULTS

We fitted cumulative Gaussian functions to the plots of percent correct against amplitude of orientation modulation and determined the threshold as the amplitude of orientation modulation giving 75% correct. The results are shown in Fig. 3. Each graph plots sensitivity to orientation modulation, defined as the reciprocal of the threshold amplitude of orientation modulation, as a function of the spatial frequency of orientation modulation. Orientation modulation functions for the SN, SQ and MF data are shown together in each graph. As can
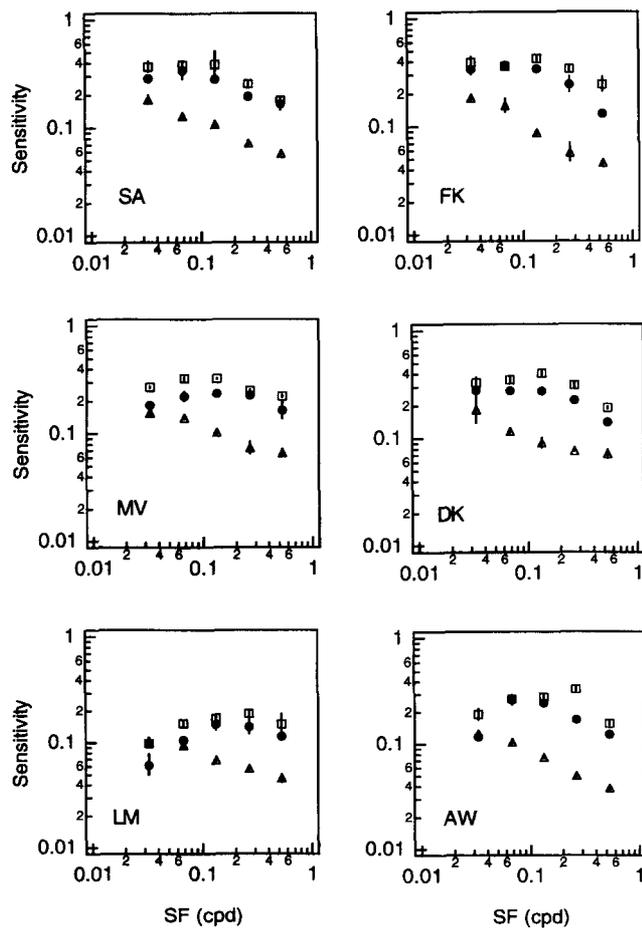
FIGURE 3. Results of experiment for six subjects. Sensitivity, defined as the reciprocal of the threshold amplitude of orientation modulation, is plotted as a function of the spatial frequency of orientation modulation. □ SQ; ● SN; △ MF. Vertical lines are upper and lower SEs of sensitivity obtained from the threshold SEs provided by the psychometric function fitting procedure.

be seen, the data in all but LM's case shows a clear hierarchy of performance: SQ > SN > MF. The SN and SQ functions are similar in shape, and are in most cases bandpass to some degree. The MF data on the other hand falls off rapidly with increasing spatial frequency of orientation modulation.

Linear Systems Analysis

Method

We have analysed our results according to three classes of linear model. For each model we used each subject's SN sensitivities to predict their individual SQ and MF data, and these predictions were made without any free parameters. We first fitted the SN data with a function which could be used to generate the different model predictions. For this we used a quadratic fit to the log transformed sensitivities when plotted as a function of log spatial frequency. The fitted function served two purposes: first to smooth the SN data for each subject and second to allow predictions for SN sensitivities at spatial frequencies other than those measured directly, which was necessary for one of the linear models tested (see below). The quadratic function generated predictions for SN sensitivities in the range of spatial frequencies from 1 to 64 times that of the lowest spatial frequency tested, which was the 0.033 c/deg condition. We found that a Gaussian fit to the SN data produced virtually identical estimates of sensitivity for those spatial frequencies not directly tested, so we concluded that our predictions were not dependent on the precise form of the function used to fit the SN data.

The predictions of the three classes of linear model were all derived by Fourier analysis using the Fast Fourier Transform (FFT) and inverse FFT routines in the graphics/analysis package IGOR (Wavemetrics Inc.) run on a Macintosh computer. FFTs were performed on each type of idealized waveform and for each spatial frequency, with the exact number of cycles of modulation that were present in the stimulus, using a resolution of 128 data points in every case. The three linear models were as follows.

(1) Single-channel peak-amplitude model. This model supposes that the peak response of a single linear operator underlies the detection of the orientation modulation in our stimuli. The predictions for this model were made by multiplying the Fourier transforms of the SQ and MF stimuli by the fitted SN sensitivities, taking the inverse Fourier transform of the result, and then measuring its peak amplitude. This is mathematically identical to taking a single linear operator whose spatial frequency tuning is given by the SN sensitivity function, convolving the operator with the SQ and MF stimuli, and measuring in each case the peak in the operator's convolution response (Bracewell, 1986, p. 110).

(2) Single-channel r.m.s. model. This model supposes that it is the r.m.s. of a single linear operator that underlies detection. The predictions for this model were made in the same way as with the single-channel peak-amplitude model described above, except that the r.m.s. of the inverse Fourier transform rather than its peak

amplitude was measured. The sequence was also performed on the SN sensitivities themselves, in order to provide SN r.m.s. values which could be used to normalize the predicted SQ and MF values.

(3) Most detectable harmonic model. This model supposes that the amplitude of the test stimulus (SQ and MF) required for detection is that required to make its most detectable harmonic reach its own independent detection threshold. To test this model we used the fitted SN sensitivities as the estimates of the detectability of the individual harmonic components in the SQ and MF stimuli. For all the subjects the most detectable harmonic in both the SQ and MF stimuli turned out to be the lowest harmonic in the stimulus. In the SQ stimulus this is the fundamental harmonic (1f) with an amplitude of $4/\pi$, or 1.273, times that of the SQ stimulus itself. The predictions for the SQ data were thus calculated by multiplying the fitted SN sensitivities at each spatial frequency by 1.273. For the MF stimulus, the lowest harmonic is the 3rd harmonic (3f), whose amplitude is $4/(3\pi)$, or 0.4244, times that of the MF stimulus itself. For the two highest spatial frequencies of the MF stimulus tested, 0.267 and 0.533 c/deg, the 3rd harmonics are at 0.8 and 1.6 c/deg respectively, and thus lie outside of the range of SN spatial frequencies actually tested (0.033–0.533 c/deg). Sensitivity to these higher-than-measured harmonics was estimated in the same way as for all the other harmonics, namely from the quadratic fit to the SN data. The predictions for the MF data were made by taking the estimates of SN sensitivity at spatial frequencies 3 times that of the MF stimulus and multiplying the result by 0.4244.

Results

The model predictions for all of the subjects are given in Fig. 4. For the single-channel peak-amplitude and most detectable harmonic models, the dotted lines through the SN data are the quadratic fits. In the case of the most detectable harmonic model the dotted line extends beyond the data to those spatial frequencies for which an estimate of sensitivity was required in order to predict the MF data (see above). For the single channel r.m.s. model, the dotted lines through the SN data represents the normalized r.m.s. values, and as expected are virtually identical to the quadratic fits. The model predictions for the SQ and MF data, which are of principal interest, are shown as solid lines. All three models correctly predict the overall ordering of sensitivity (SQ > SN > MF), as well as the overall shape of the functions. Note in particular how the very differently shaped SN functions in DK's and LM's data, representing respectively the two extremes in the range from lowpass to markedly bandpass, is each able to predict much of the individual subject's SQ and MF data. Before providing a more quantitative comparison of the three classes of linear model however, it is necessary to consider the possible reasons for one noticeable failure of all three models: to predict the performance at the two highest spatial frequencies of the MF stimulus.

Effects of Undersampling and Orientation Noise

Performance at the two highest spatial frequencies of the MF stimulus is consistently superior to the predictions of any of the linear models. The fact that performance is higher than expected strongly suggests that lower harmonic components (in the orientation domain) were introduced into the neural image through aliasing, as a result of the manner in which the stimuli were physically constructed and because of additional limitations imposed by visual processing. There are two likely causes of aliased frequencies in our stimuli: undersampling and the effects of noise in the representation of local orientation. Undersampling could arise from a combination of three factors. First, because the stimuli were generated by discretely sampling the waveform by a finite (1000) number of Gabor patches; second, because the positions of those Gabor patches were completely random in the stimulus; third, because the visual system may only have effectively sampled the orientation information from a proportion of the Gabor patches present. Orientation noise was introduced into the stimuli at generation (the orientation of each Gabor patch was drawn randomly from a Gaussian distribution with a mean defined by the point on the waveform and a SD of 10 deg), and the visual system would undoubtedly add further noise during the process of transduction.

To examine the likely effects of both undersampling and orientation noise we carried out the following simulations. First, we produced a one-dimensional representation of each waveform in an 8192 point array, and then sampled the waveform with S points, whose positions within the array were randomized. S was a maximum of 1000 (the number of Gabor patches in each stimulus). The random positioning of S points in a much larger array captured both the discrete nature of the sampling and the randomization of the sample positions. We then added noise to each sampled point in the model waveform by an amount drawn randomly from a Gaussian distribution with a SD given by σ_n/T. σ_n was the SD of the added orientation noise, scaled by $1/T$ to enable the calculations to be performed on waveforms with unit, rather than threshold, amplitude. The sampled waveform with added orientation noise was then subject to the same linear modeling as described in the previous section, with the difference that each predicted point was the mean of 10 simulations, each carried out with a freshly generated set of S random sample points. Figure 5(a) shows the effect of S (sample size), on the predictions of the single channel r.m.s. model for subject LM. The data in Fig. 5(a) has been normalized to the completely sampled, noise free, sine-wave condition. Figure 5(b) shows the effect of adding different levels of σ_n (orientation noise) to the $S = 1000$ sampled condition. Figure 5(a, b) demonstrates how *in principle* either a significant amount of undersampling, or a significant amount of orientation noise, can predict the relative increase in the performance of the MF high spatial frequency data, and that these two factors have little effect on the other stimulus conditions.
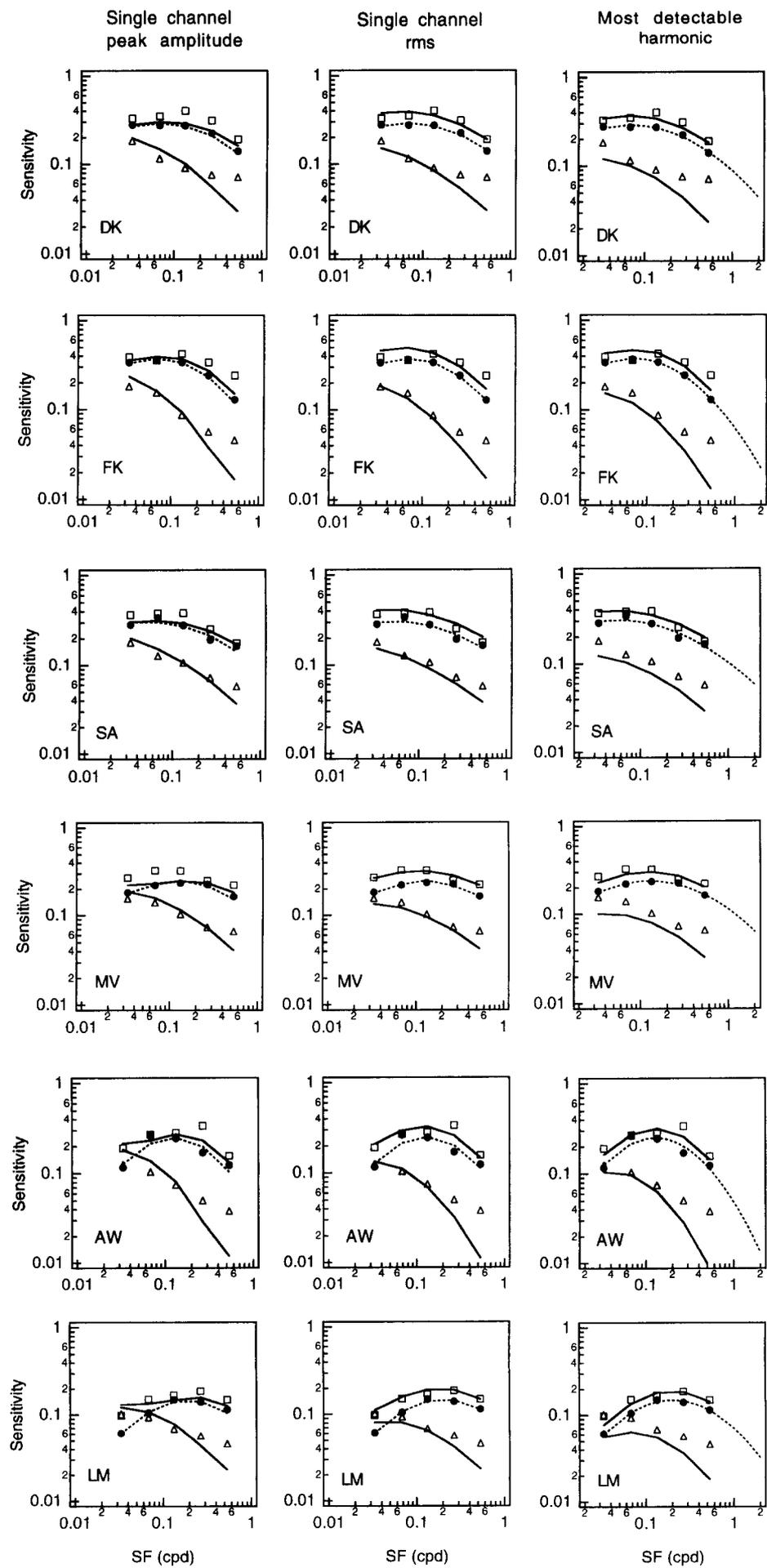
FIGURE 4. *Caption on facing page.*
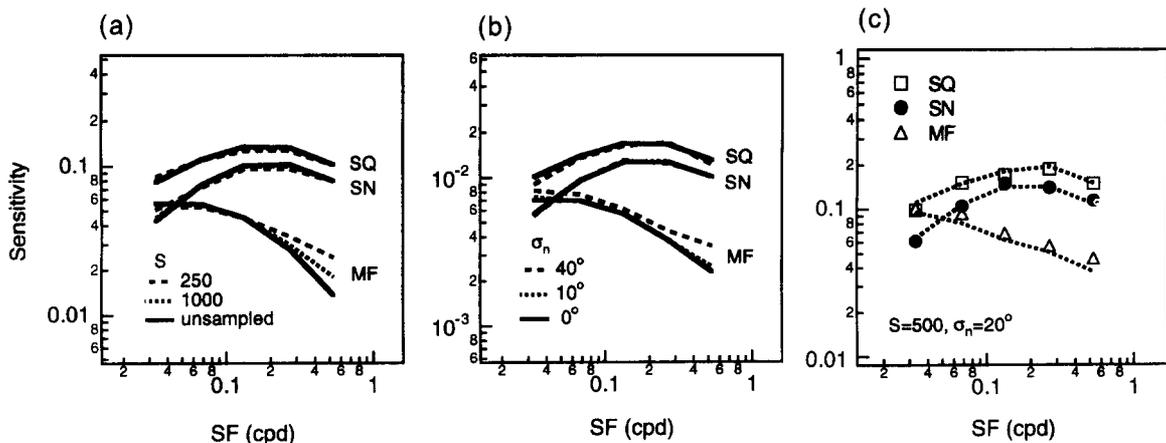
FIGURE 5. Effects of (a) undersampling, (b) added orientation noise and (c) both, on subject LM's results using the single-channel r.m.s. model. In (a) the three conditions are: unsampled, 1000 sample points and 250 sample points. The curves have all been normalized to the unsampled SN condition and the actual subject's data points [shown in Fig. 5(c)] have been omitted to minimize cluttering. The effect of reducing the number of samples is to selectively elevate sensitivity to the high spatial frequency MF stimuli. In (b) the consequence of adding orientation noise, σ_n, to the 1000 sample condition shows a similar pattern. (c) The prediction of LM's results using the single-channel r.m.s. model for a simulation involving 500 sample points and a value of σ_n of 20 deg. For further details see text.

Finally Fig. 5(c) shows a single prediction to demonstrate the most likely scenario that *both* undersampling and orientation noise contribute to the improved performance in the high spatial frequency MF data. Here we have arbitrarily set S to 500 (half the actual sample size) and σ_n to 20 (twice the level of introduced orientation noise).

Although these model predictions strongly argue for aliasing as the cause of the high spatial frequency MF performance, there may of course be other contributory factors such as a distortion of the shape of the waveform resulting from non-linear grouping processes (see Discussion). Given such a possibility, and given our desire to apply the linear models to the data without any free parameters, we have taken the conservative approach of simply removing the two high spatial frequency MF data points for the quantitative comparison of the three linear models now described.

Quantitative Comparison of the Three Linear Models

An overall picture of the relative predictive power of the three classes of linear model is given in Table 1. Table 1 shows χ^2 values for the three model predictions when χ^2 was calculated thus:

$$\chi^2 = \frac{1}{n}\sum_{i=1}^{n}[(t_i - p_i)/\sigma_i]^2 \tag{2}$$

where t_i is the measured threshold of orientation modulation for the ith spatial frequency, σ_i the ith SE of the threshold estimated from the psychometric function, p_i

TABLE 1. χ^2 values for model predictions

	Single-channel peak-amplitude		Single-channel r.m.s.		Most detectable harmonic	
	SQ	MF	SQ	MF	SQ	MF
FK	9.2	14.3	7.9	1.6	6.2	15.7
MV	691.0	21.3	8.2	26.9	52.0	236.8
SA	25403	2.4	2451	3.4	227	21.2
AW	14.1	22.9	5.5	4.3	5.6	18.1
LM	6.3	6.8	1.6	9.0	4.6	78.5
DK	11.9	14.2	2.1	0.83	2.4	8.0

For the MF predictions only data for the three lowest spatial frequencies (0.033–0.133 c/deg) are included.

the ith model prediction and n the number of data points per model fit, which was 5 for the SQ condition and 3 for the MF condition (the 0.267 and 0.533 c/deg MF conditions were left out for reasons given in the previous section). The smaller the value of χ^2, the better the fit to the data, though it should be noted that the values are generally quite high owing to the very tiny values of σ (see the error bars on Fig. 3). A good measure of the relative fits of the three models is the ratio of χ^2 values computed for each subject for a given pair of the linear models. If we take the (geometric) mean ratio of χ^2 across subjects and across the two stimulus conditions (SQ and MF), we find that the single- channel r.m.s. model produces one-quarter the value compared to the single-channel peak-amplitude model, and just over one-third the value compared to the most detectable harmonic model.

FIGURE 4 *(opposite)*. Predictions for the three linear models. The dashed line through the SN data is the best fitting quadratic for the single-channel peak-amplitude and the most detectable harmonic models, and the normalized r.m.s. values for the single-channel r.m.s. model. In the case of the most detectable harmonic model, the dashed line continues beyond the subjects' data to include those spatial frequencies for which an estimate of SN sensitivity was required in order to predict the two highest spatial frequency MF data points. Continuous lines through the SQ and MF data are the individual model predictions generated from each subject's SN data.
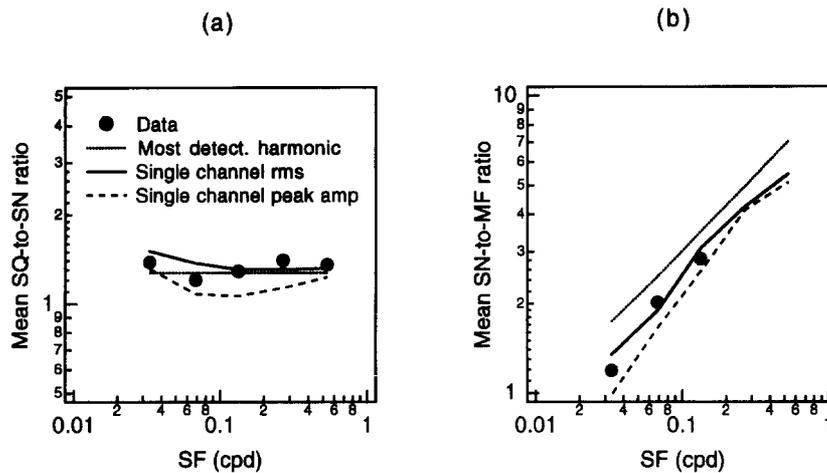
FIGURE 6. Ratio of (a) SN-to-SQ sensitivity, (b) SN-to-MF sensitivity, averaged across all subjects. In (b) only the three lowest spatial frequencies are compared. The lines through the data are the three model predictions averaged across subjects.

A different method of comparing the relative predictive power of the three linear models is illustrated in Fig. 6. Figure 6(a) shows the average SN-to-SQ sensitivity ratio and Fig. 6(b) the average SN-to-MF sensitivity ratio across all six subjects. Each graph also shows the predictions from the three models, again averaged across all six subjects. In Fig. 6(a), the closeness of the predictions of the single-channel r.m.s. and most detectable harmonic models is apparent. Figure 6(b) demonstrates that, at least for the first three data points included in the quantitative comparisons, the single-channel r.m.s. model gives the best fit, albeit only marginally. In conclusion the quantitative comparison of the three classes of linear model suggests that the single-channel model gives the best fit to both the SQ and MF data alike.

DISCUSSION

In this study we have shown that the detection of orientation modulation can be adequately modelled by a single linear mechanism. We found that the measured sensitivities for detecting the orientation modulation in SN modulated textures were able to predict to a surprising degree of accuracy sensitivity to both SQ and MF modulated textures in all the subjects tested. The linear model which gave the best fit to both the SQ and MF data consisted of a single linear channel whose r.m.s. response determined detection. Two other linear models performed slightly less well: a single linear mechanism whose peak amplitude response determined performance, and one in which performance was determined by sensitivity to the most detectable harmonic in the stimulus.

The r.m.s. of the physical waveform

Given how well the single-channel r.m.s. model performs, it might be supposed that a model based simply on the r.m.s. of the physical waveform itself, rather than that of the output of a putative mechanism in response to that waveform, might also give a good account of the

data. The r.m.s.'s of unit amplitude SQ, SN and MF waveforms are 1.414, 1 and 0.414 respectively when normalized to the SN waveform, and this clearly predicts the overall ordering of sensitivity we observed in our data (with the notable exception of subject LM's low spatial frequency results). However, a model based simply on the physical r.m.s. of the stimuli, or indeed the mean absolute deviation, would predict flat orientation modulation functions for all stimuli. A modified version of a physically-based r.m.s. model incorporating the attenuation characteristics described by the SN data would give non-flat functions, but they would nevertheless be parallel, unlike what was found. A physically-based r.m.s. model is thus untenable.

Single- vs multiple-channel models

Linear systems analysis has been remarkably successful in modelling threshold luminance vision. It was the comparison of the detectability of SN, SQ and MF stimuli in the luminance domain that led Fergus Campbell and his colleagues to conclude that multiple independent channels underlie the detection of luminance patterns (Campbell & Robson, 1968; Campbell et al., 1971, 1978). They found that multi-harmonic luminance patterns were detected once their most detectable harmonic component reached its own independent threshold. This is formally identical to the most detectable harmonic model tested here, which we found to give a good fit to the data but not as good as that provided by the single-channel r.m.s. model. Although our results offer the possibility that the detection of orientation modulation, unlike luminance detection, is mediated by a single- rather than multiple-channel mechanism, at least at a given spatial scale of luminance, one must be cautious for two reasons. First, the single-channel r.m.s. model predictions are only marginally superior to those of the most detectable harmonic model. Second, we cannot rule out the possibility that there may have been some probability summation between the lowest and most detectable, and the higher, harmonics in the SQ and MF stimuli, even though the higher

harmonics were of significantly lower amplitude. Our single-channel r.m.s. model could thus constitute an emergent property of multiple-channel texture mechanisms. Further experiments will be needed to categorically distinguish between single- and multiple-channel accounts of our data. One such experiment would be to test whether textures consisting of one or other of two equally detectable harmonic components, widely separated in spatial frequency, are discriminable at threshold. The single-channel r.m.s. model predicts that they should not be discriminable, whereas a multiple-channel model predicts that they should be discriminable.

Nature of first-stage input

Our model operator constitutes a description of the putative "second-stage" of orientation gradient detection, which could in principle integrate the outputs of an array of "first-stage" filters, such as simple or complex cells, which process the luminance detail of the micropatterns in each stimulus. What do our results tell us about the nature of this first-stage input? First, one must be clear about what our results do *not* tell us. They do not tell us anything about the kind of point non-linearities which might be imposed prior to, at the stage of, or immediately after the processing of the luminance detail in our stimuli by the first-stage filters. This issue has been addressed in a number of studies on texture segregation and many models now include point non-linearities such as half-wave rectification, full-wave rectification and contrast compression imposed on the early filter outputs (Graham et al., 1992; Malik & Perona, 1990; Rubenstein & Sagi, 1990; Gorea & Papathomas, 1993). Can we nevertheless say anything at all about the nature of its input? The good data fits provided by the linear models are consistent with the input being a more-or-less veridical (though quite possibly noisier than in the stimulus itself) map of local orientations. This might appear at odds with the significant body of research supporting the existence of facilitatory and inhibitory interactions between neighbouring oriented filters. These non-linear interactions are believed to emphasize local features which lie along straight lines and smooth curves, that is features which display "good continuity" (Zucker & Davis, 1988; Link & Zucker, 1988; Field, Hayes & Hess, 1993). Such interactions exploit the fact that good continuity is a common feature of contours and textures in natural scenes (Parent & Zucker, 1989; Field et al., 1993). There is psychophysical evidence for the involvement of curvilinear gouping processes in orientation-defined textures (Zucker, 1983, 1985; Link & Zucker, 1988; Or & Zucker, 1989) and recently a number of neural network models have been developed to implement their operation (Lowe, 1988; Parent & Zucker, 1989; Sha'ashua & Ullman, 1988; Gigus & Malik, 1991). That these processes may be implicated in the detection of the orientation modulation in our displays can be seen by the strong subjective flow patterns observed in the stimuli shown in Fig. 2. Zucker (1985) has defined a flow pattern as "a dense covering of a surface with a family of curves that are locally parallel

almost everywhere". In our previous paper (Kingdom et al., 1995) we provided some evidence which was consistent with such non-linear grouping processes. We found a non-linear relationship between the squared detection thresholds for SN orientation modulation and the amount of added orientation noise variance. We argued that this was inconsistent with a linear statistical decision model. Instead, increasing the amount of orientation noise would increasingly disrupt the possibility of linking between neighbouring micropattern detectors with similar orientation preferences and would degrade performance over and above that which would be expected solely on statistical grounds. We must conclude therefore that while non-linear grouping processes are undoubtedly involved in the processing of our stimuli, they probably act primarily to sharpen the signal rather than distort its shape.

Receptive field organization of second-stage

The account of our data in terms of a putative linear operator has at first sight much in common with a number of recent models which emphasize the extraction of local texture gradients as the basis for image segmentation (e.g. Malik & Perona, 1990; Rubenstein & Sagi, 1990; Landy & Bergen, 1991; Nothdurft, 1991). Our SN orientation modulation functions are in most cases bandpass to some degree, as we found in our previous study (Kingdom et al., 1995), consistent with the operation of orientation gradient detectors. The important point about our SN data however, is that it shows maximal sensitivity to relatively low spatial frequencies of orientation modulation. Peak sensitivity occurred at an average spatial frequency of orientation modulation of 0.093 c/deg, which is consistent with our previous study, where the figure was 0.09 c/deg. This figure was calculated by taking the geometric average of the peak spatial frequencies estimated from the quadratic fits to the SN data. A figure of about 0.09 c/deg implies that peak sensitivity to orientation modulation occurs for textures modulated at a spatial frequency more than 50 times lower than the dominant luminance spatial frequency of the micropatterns, which was 4.8 c/deg. The receptive field profile of the linear operator which can predict performance with our stimuli can be obtained from the inverse Fourier transform of the SN data. The results are shown in Fig. 7 for the real part of the inverse Fourier transform, which generates an even-symmetric filter profile. In general the receptive fields have narrow centres with broad but shallow inhibitory surrounds. Each plot in Fig. 7 gives W_c, the width of each receptive field centre (defined as the distance between the two zero-crossings on either side of the centre). In a recent survey by Gurnsey and Laundry (1992) of the Malik and Perona (1990) and Rubenstein and Sagi (1990) texture-segregation models, which were inspired by human psychophysical data, Gurnsey and Laundry concluded that they employed second-stage filters whose receptive fields covered an area of about 3×3, or 9 in total, micropatterns. In our stimuli mean micropattern density was 4.8 micropatterns/deg^2, so a 3 deg diameter square
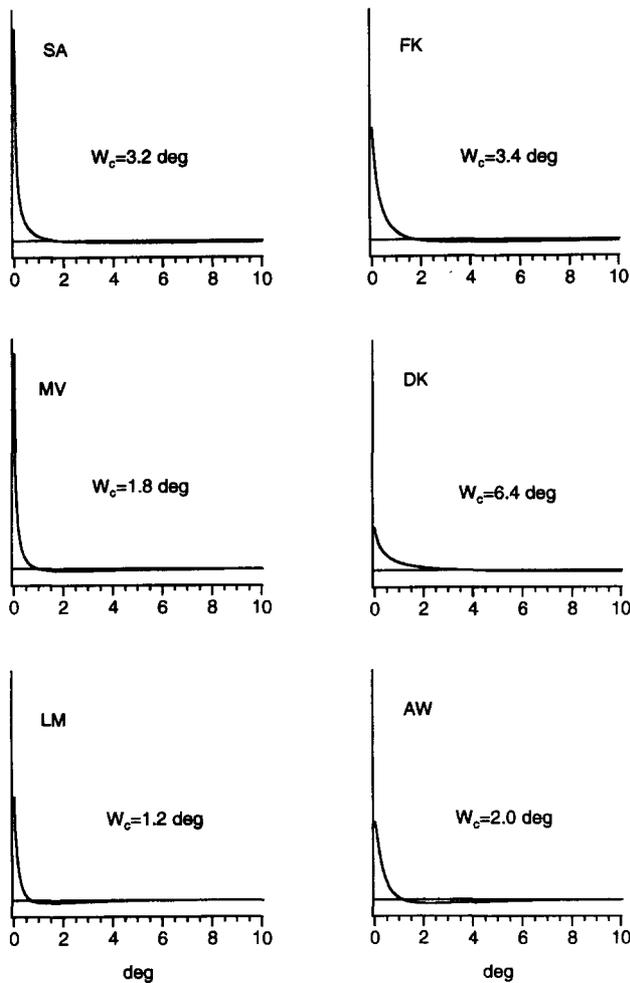
FIGURE 7. Each plot shows one half of the even-symmetric receptive field profile of the putative operator underlying performance. Each plot is derived from the real part of the inverse Fourier transform of each individual subject's data. W_c gives the estimated width of the centre of the receptive field, defined as the distance between the two zero-crossings on either side of the centre.

receptive field centre (the mean value of W_c from Fig. 7) would *alone* cover about 6.6 × 6.6, or 43 in total, micropatterns. It is not realistic to give a figure for the number of micropatterns covered by the receptive field surround because it is so shallow, but it is clearly an order of magnitude greater than the centre. More recently Rubenstein and Sagi (1993) have measured the detectability of square-wave texture modulations (their square-waves had a fixed peak-to-trough amplitude of 90 deg, and performance was limited by a post-stimulus mask presented at a variable inter-stimulus-interval) and on the basis of the results proposed second-stage filters whose centres spanned about seven micropatterns, and whose surrounds extended 2–3 times that of the centres. This is remarkably similar in size to our model second-stage operators. An arguably better way of making the comparison is in terms of the number of cycles of the Gabor micropattern carrier across the receptive field centre, rather than number of micropatterns, for reasons given below. Our micropatterns were 4.8 c/deg, and thus a 3 deg wide receptive field centre would cover about 14 carrier cycles. Rubenstein and Sagi proposed a 7 deg

wide receptive field centre, which for their 2 c/deg micropatterns used in their orientation gradient task also implies 14 carrier cycles. There are many differences between the stimuli and tasks used in our study and those others we have cited, and these could account for many of the differences in the proposed size of the texture operators. One must conclude however that for the detection of orientation gradients, our data, together with the recent study of Rubenstein and Sagi (1993), argue for second-stage filters with larger receptive fields than has previously been supposed. This conclusion is also in keeping with the more "region-based", rather than "edge-based" approach to texture segregation suggested by Gurnsey and Laundry (1992), who found that blurring or occluding the sharp borders in textures did not substantially impair segregation performance. A caveat to this argument however is that the large receptive field operators that we are suggesting may be too coarse for *locating* texture boundaries, and since texture boundaries can presumably be located with reasonable accuracy, it may be that different mechanisms subserve their localisation as opposed to their detection.

What might be the physiological basis of our second-stage operator? Earlier we made the reservation that the single-channel r.m.s. model might represent an emergent property of multiple-channel texture mechanisms. Whether single- or multiple-channel however, our putative operator nominally integrates all the orientation information that lies within its receptive field, and gives a signal dependent on both the magnitude and the spatial frequency of the orientation *modulation* it samples. It is therefore tuned to orientation *differences* rather than to orientation *per se*. One way in which this could be implemented physiologically would be to have a centre mechanism which summed the activity of first-stage filters tuned to one orientation, and a surround mechanism which summed the activity of first-stage filters tuned to the orthogonal orientation, as in the recent model of Rubenstein and Sagi (1993). Some supportive neurophysiological evidence for such a receptive field organization comes from single unit recordings in monkey V1 (Knierim & Van Essen, 1992), and MT (Olavarria, De Yoe, Knierim, Fox & Van Essen, 1992).

Although we have demonstrated that a single-channel model provides a good description of the second-stage of orientation-defined texture processing, it is very likely that different sized second-stage filters exist for processing different spatial scales of their first-stage inputs. Kingdom *et al.* (1995) found that the SN orientation modulation function was *scale invariant*, i.e. it did not shift significantly in its position on the spatial frequency axis with a change in viewing distance, if spatial frequency was measured in stimulus rather than retinal units (e.g. c/cm rather than c/deg). We argued, following Bergen (1991), that this could result from orientation gradient operators being tied to micropattern size or spatial frequency, a feature incidentally incorporated in some models of texture processing (e.g. Voorhees & Poggio, 1988; Sagi, 1990). Whether there are indeed different orientation gradient operators for processing

different spatial scales of luminance contrast is, however, a matter requiring more direct empirical verification than the demonstration of scale invariance.

Abrupt vs smooth texture gradients

Finally, we return to the question we first posed in the Introduction. Are the mechanisms for detecting abrupt texture gradients the same as those for detecting smoothly varying texture gradients? We have shown that the detectability of stimuli with smooth spatial variations in orientation (the SN stimuli) can predict the detectability of stimuli with abrupt (the SQ stimulus), and with both abrupt and smooth (the MF stimulus), spatial variations in orientation. We conclude therefore that abrupt and smooth spatial variations in orientation-defined textures are detected by a common mechanism which mimics the operation of a single linear operator.

REFERENCES

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

Beck, J. (1982). Textural segmentation. In Beck, J. (Ed.), *Organization and representation in perception.* London: Lawrence Erlbaum.

Bergen, J. R. (1991). Theories of visual texture perception. In Regan, D. (Ed.), *Vision and visual dysfunction* (Vol. 10B). New York: Macmillan.

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

Blake, A., Bulthoff, H. H. & Sheinberg, D. (1991). Shape from texture: Ideal observers and human psychophysics. *Vision Research, 33,* 1723–1737.

Bracewell, R. N. (1986). *The Fourier transform and its applications.* New York: McGraw-Hill.

Buckley, D. & Frisby, J. (1993). Interaction of stereo, texture and outline cues in the shape perception of three-dimensional ridges. *Vision Research, 33,* 919–934.

Campbell, F. W. & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, London, 197,* 551–566.

Campbell, F. W., Howell, E. R. & Johnstone, J. R. (1978). A comparison of threshold and suprathreshold appearance of gratings with components in the low and high spatial frequency range. *Journal of Physiology, 284,* 193–201.

Campbell, F. W., Howell, E. R. & Robson, J. G. (1971). The appearance of gratings with and without the fundamental Fourier component. *Journal of Physiology, London, 217,* 17–19.

Cutting, J. & Millard, R. (1984). Three gradients and the perception of flat and curved surfaces. *Journal of Experimental Psychology, General, 113,* 198–216.

Davis, L., Janos, L. & Dunn, S. M. (1983). Efficient recovery of shape from texture. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 5,* 485–492.

Field, D. J., Hayes, A. & Hess, R. F. (1993). Good continuation and the association field: Evidence for local feature integration in the visual system. *Vision Research, 32,* 173–193.

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

Gigus, Z. & Malik, J. (1991). Detecting curvilinear structure in images. Technical report UCB/CSD 91/619, Computer Science Division (EECS) University of California at Berkeley, Calif.

Gorea, A. & Papathomas, T. V. (1993). Double opponency as a generalized concept in texture segregation illustrated with stimuli defined by color, luminance, and orientation. *Journal of the Optical Society of America A, 10,* 1450–1462.

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

Gurnsey, R. & Browse, R. (1988). Aspects of visual texture discrimination. In Pylyshyn, Z. (Ed.), *Computational processes in human vision: An interdisciplinary approach.* Norwood, N.J.: Ablex.

Gurnsey, R. & Laundry, D. S. (1992). Texture discrimination with and without abrupt texture gradients. *Canadian Journal of Psychology, 46,* 306–332.

Julesz, B. (1975). Experiments in the visual perception of texture. *Scientific American, 232,* 34–43.

Kanato, L. & Chou, T. (1989). Shape from texture: General principle. *Artificial Intelligence, 38,* 1–48.

Kingdom, F. A. A. & Keeble, D. R. T. (1994). The detection of "edges" in orientation-defined textures. *Investigative Ophthalmology and Visual Science (Suppl.), 34,* A4194, p. 2161.

Kingdom, F. A. A., Keeble, D. R. T. & Moulden, B. (1955). Sensitivity to orientation modulation in micropattern-based textures. *Vision Research, 35,* 79–91.

Knierim, J. J. & Van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neuroscience, 67,* 961–980.

Lamme, V. A., Van Dijk, B. W. & Spekreijse, H. (1992). Texture segregation is processed by primary visual cortex in man and monkey. Evidence from VEP experiments. *Vision Research, 32,* 797–807.

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

Linderberg, T. & Garding, J. (1993). Shape from texture from a multi-scale perspective. *IEEE Proceedings of the Fourth International Conference on Computer Vision, 2,* 683–691.

Link, N. K. & Zucker, S. W. (1988). Corner detection in curvilinear dot grouping. *Biological Cybernetics, 59,* 247–256.

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

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

Olavarria, J. F., DeYoe, E. A., Knierim, J. J., Fox, J. M. & Van Essen, D. C. (1992). Neural responses to visual texture patterns in middle temporal area of the macaque monkey. *Journal of Neuroscience, 68,* 164–181.

Or, U. H. & Zucker, S. W. (1989). Texture fields and texture flows: Sensitivity to differences. *Spatial Vision, 4,* 131–139.

Parent, P. & Zucker, S. W. (1989). Trace inference, curvature consistency, and curve detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11,* 823–839.

Rubenstein, B. S. & Sagi, D. (1990). Spatial variability as a limiting factor in texture-discrimination tasks: Implications for performance asymmetries. *Journal of the Optical Society of America, 7,* 1632–1643.

Rubenstein, B. S. & Sagi, D. (1993). Effects of foreground scale in texture discrimination tasks: Performance is size, shape, and content specific. *Spatial Vision, 7,* 293–310.

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

Sha'shua, A. & Ullman, S. (1988). Structural saliency: The detection of globally salient structures using a locally connected network. *IEEE Proceedings of the Second International Conference on Computer Vision,* 321–327.

Stevens, K. & Brookes, A. (1988). Integrating stereopsis with monocular interpretations of planar surfaces. *Vision Research, 28,* 371–386.

Todd, J. & Akerstrom, R. (1987). Perception of three-dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 242–255.

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

Witkin, A. (1981). Recovering surface shape and orientation from texture. *Artificial Intelligence, 17,* 17–45.

Zucker, S. W. (1983). Computational and psychophysical experiments in grouping: Early orientation selection. In Beck, J., Hope, B. & Rosenfield, A. (Eds), *Human and machine vision.* London: Academic Press.

Zucker, S. W. (1985). Early orientation selection: Tangent fields and the dimensionality of their support. *Computer Vision, Graphics and Image Processing, 32*, 73–103.

Zucker, S. W. & Davis, S. (1988). Points and end-points: A size/spacing constraint for dot grouping. *Perception, 17*, 229–247.