Artificial neural networks simulating visual texture segmentation and target detection in line-element images

ANDREW J. SCHOFIELD* AND DAVID H. FOSTER‡

Department of Communication and Neuroscience, Keele University, Staffordshire ST5 5BG, U.K.

SUMMARY

Measurements were made of human observers' performance both in segmenting regions of line-elements and in detecting line-element targets in stimuli containing several orientations. Performance was modelled by four artificial neural networks constructed from processing units trained to mimic the gross functionality of certain loosely defined classes of cortical cells. Model 1 contained modules sensitive to absolute orientation only, and it provided a poor fit to the human-performance data. Model 2 contained modules sensitive to orientation contrast: the outputs of these modules could be suppressed with fields of uniformly oriented line-elements. Model 3 contained orientation-contrast-sensitive modules of a different type: their outputs could be suppressed with fields of randomly oriented line-elements. Models 2 and 3 both successfully processed line-element arrays with orientation heterogeneities, but these models still provided inadequate fits to the human-performance data. Model 4 contained both types of orientation-contrast-sensitive modules; this model was able to account for human performance in the segmentation and detection tasks, both qualitatively and quantitatively.

1. INTRODUCTION

As part of the basic process of detecting and discriminating objects in complex scenes, the image presented to the eye is visually divided or segmented into a foreground that is of immediate interest and a background that is not. The segmentation of regions in a scene is performed speedily and effortlessly, and has been attributed to the early stages of visual processing, often referred to as pre-attentive (Julesz 1962, 1981*a*) or as involving distributed attention (Beck 1972). The detection of a small object (or target) in a scene – a task related to image segmentation – has also been assumed to be determined by preattentive visual processes (Neisser 1967; Treisman 1977; Javadnia & Ruddock 1988; Foster & Ward 1991*a*). A variety of psychophysical studies have shown that images can be segmented on the basis of variations in luminance, colour, motion, depth (stereo disparity), and texture (Julesz 1962, 1980, 1981*a*, *b*, 1984; Beck 1966, 1972, 1983; Olson & Attneave 1970; Nakayama *et al.* 1985; Nothdurft 1985*a*, *b*, 1991, 1993; Landy & Bergen 1991). The work reported here concentrated on textures composed of oriented line-elements. The aim was to compare human performance in texture-segmentation and line-target-detection tasks with models of that performance based on artificial neural networks. (A preliminary report of this work has been published already; see Schofield & Foster 1993.)

* Present address: Metropolitan Police Forensic Science Laboratory, 109 Lambeth Road, London, SE1 7LP, U.K.

‡ Present address: Department of Vision Sciences, Aston University, Aston Triangle, Birmingham B4 7ET, U.K.

(*a*) *Textures of oriented lines*

According to the texton theory of texture segmentation (Julesz 1981*a*, 1984), two or more textures can be visually discriminated if the underlying pattern of elements contains different numbers of elementary cues or 'textons', of which three types were proposed: the orientation of elongated shapes (lines, bars, ellipses), line crossings, and line terminations. Line crossings and terminations may give rise to luminance artefacts that aid segmentation (Nothdurft 1990); when these artefacts are eliminated, only orientation is left as a salient cue for image segmentation, a result which is consistent with the finding that the orientations of the component parts of texture elements (lines, bars, ellipses) are generally more important for segmentation than the relative spatial arrangements of those parts (see also Beck 1966, 1972; Olson & Attneave 1970; Foster & Mason 1980). (Relative spatial arrangement can, of course, be important if texture density is not constant, and, in isotropic textures, variations in local spatial-frequency content can provide a cue for segmentation.) Figure 1*a* shows an example of a line-element array containing two regions that are segmentable by virtue of the differences in the orientations of the line-elements making up the two regions.

(*b*) *Orientation contrast*

Performance in segmenting textures of oriented line-elements clearly depends on the difference in orientations between the two texture regions; that is, on the 'orientation gradient' or 'orientation contrast'

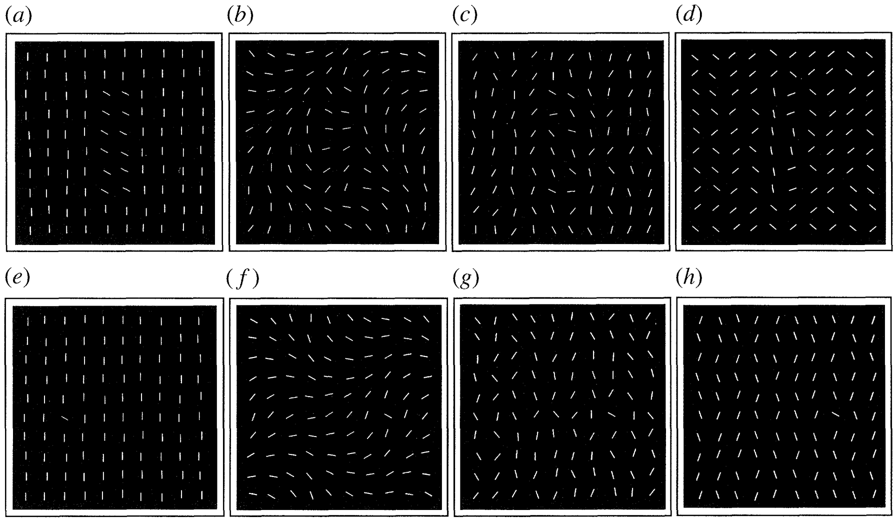

Figure 1. Example displays. Top row: texture-segmentation displays with (*a*) no orientation heterogeneities; (*b*) progressive orientation variations; (*c*) random orientation variations; (*d*) grouped orientation variations. Bottom row: target detection displays with (*e*) no orientation heterogeneities; (*f*) progressive orientation variations; (*g*) random orientation variations; (*h*) grouped orientation variations.

(Nothdurft 1985*b*; Landy & Bergen 1991). (To emphasise the difference between line-element orientation and orientation contrast, the former – defined implicitly with respect to some reference frame – is referred to as 'absolute orientation'.) In practice, both orientation contrast and absolute orientation are important in texture-segmentation and line-target-detection tasks (see, for example, Olson & Attneave 1970; Beck 1973, 1982; Treisman & Souther 1985; Gurnsey & Browse 1987; Rubenstein & Sagi 1990; Nothdurft 1991; Foster & Westland 1995); the effects of absolute orientation are revealed as marked orientational asymmetries in performance (Treisman 1985; Treisman & Gormican 1988; Foster & Ward 1991*a*, *b*; Marendaz *et al.* 1991, 1993; Meigen *et al.* 1994).

It is possible to reduce or eliminate the effects of absolute orientation in some texture-segmentation tasks. In figure 1*b* line-element orientation varies progressively over the display so that the two regions contain line-elements at various orientations; moreover, orientations within the central region also occur in the background region (see Nothdurft 1991). The two regions are visually segmentable because the difference in line-element orientations across the boundary of the regions is greater than the difference in line-element orientations across a similar distance within the regions. Figures 1*c* and *d* show two more examples of displays in which the effects of absolute orientation have been reduced by introducing orientation heterogeneities, here non-progressive changes in line-element orientation.

(*c*) *Target detection*

Detecting a single target element in a background of other elements can be regarded as a limiting case of texture segmentation. Although not all cues serve segmentation and target detection equally well (Wolfe 1992), it seems that orientation (Treisman & Gormican 1988) and orientation contrast (Nothdurft 1991, 1993) do have similar roles in the two kinds of tasks. Figure 1*e* shows a target element defined by virtue of its unique orientation, and figure 1*f* a target element defined by virtue of its local orientation contrast. Figures 1*g* and *h* show examples of displays in which the effects of absolute orientation have again been reduced by introducing non-progressive orientation heterogeneities.

(*d*) *Neurophysiological data*

Many cells in area V1 of primate visual cortex respond selectively to moving lines (and moving bars and edges) of appropriate orientation, speed, and direction of movement. The variations of responses of cells in area V1 with respect to stimulus orientation have been well documented (Hubel & Wiesel 1968; Schiller *et al.* 1976; Gilbert 1977; De Valois *et al.* 1982; Dean & Tolhurst 1983). Orientation-selective cells divide typically into two broad categories: simple cells, sensitive to both the orientation of lines and their position within the cell's receptive field; and complex cells, sensitive to the orientation of lines but not to their exact position. Cells that are sensitive to line length have also been detected; they have been referred to as end-stopped and can be classified as simple or complex in the dependence of their responses on stimulus position (Dreher 1972). The class of complex cells contains many divisions of which the standard-complex cell is most like simple cells in its behaviour and location within the cortex (Gilbert 1977). Simple and standard-complex cells have a range of orientation-tuning bandwidths from about 15° to 100° (full width at half height) and a range of preferred orientations.

Multiple-line-element displays have been used as stimuli to cells in area V1 in several studies (Van Essen *et al.* 1989; Knierim 1991; Knierim & Van Essen 1992). In those studies, many cells have been found whose response to a single line-element was modified when that line-element was surrounded by other line-elements lying outside the classically defined receptive field. These orientation-contrast-sensitive cells have a variety of response characteristics, three of which are of interest here. Notice, however, that although cells may be classified according to these characteristics, the classification may not represent a natural functional division (Knierim & Van Essen 1992). Thus all of the cells responded strongly to patterns of lines with orientation contrast, but true 'orientation-contrast' cells responded strongly to a single line-element at the preferred orientation and with equal strength to the same (central) element surrounded by line-elements oriented orthogonally (orientation-contrast stimuli). This response was suppressed when the central line-element was surrounded by line-elements at the same orientation (uniform-orientation stimuli), or by randomly oriented line-elements (random-orientation stimuli). 'Uniform-suppressed orientation-contrast' cells responded strongly to single line-elements, to orientation-contrast stimuli, and to random-orientation stimuli, but more weakly to uniform-orientation stimuli. 'Random-suppressed orientation-contrast' cells responded strongly to single line-elements, to orientation-contrast stimuli, and to uniform-orientation stimuli, but more weakly to random-orientation stimuli. Many cells in each response group had no preferred orientation (responding well to line-elements at all orientations), yet were sensitive to orientation-contrast stimuli (Knierim & Van Essen 1992).

Models incorporating true orientation-contrast units alone do not perform well in texture-segmentation tasks (Schofield 1993). These simpler models failed to segment images whenever the orientation heterogeneities exceeded the bandwidth of the majority of orientation-sensitive units (i.e. 15°): this study therefore concentrated on random-suppressed and uniform-suppressed units.

(*e*) *Artificial neural networks*

The artificial neural networks (ANNS) used here were multilayer perceptrons trained under the back-propagation rule (Rumelhart *et al.* 1986). These networks consist of several processing elements joined to one another and to the network inputs by weighted connections. In response to the presentation of an input to the network, each processing unit calculates the weighted sum of its inputs and (in this study) applies either a sigmoid or a linear output function. Processing units are arranged in interconnected layers. Processing starts at the input layer and continues forward through the network until the output values of the units in the output layer have been calculated. During training, the values of the units in the output layer are compared with a set of desired values and an error signal for each output unit is calculated. These errors are then back-propagated to the preceding layers in the network until error signals have been calculated for all units in the network. The values of the weights are then updated so as to reduce the error signals. This procedure is repeated so that the training images are each presented many times. Once training is complete and the network has learned to produce the desired outputs for each training image, error back-propagation and weight-updating cease.

Although networks trained by back-propagation cannot model the development of biological cells (Rumelhart *et al.* 1986; Moorhead *et al.* 1989), they have been used successfully to model some aspects of their function (Lehky & Sejnowski 1988, 1990; Moorhead *et al.* 1989). The ANNS considered here should be distinguished from those essentially 'hard-wired' neural-network models that, *ab initio*, have processing units and connections with specific neurophysiologically relevant properties (e.g. Fogel & Sagi 1989; Malik & Perona 1990; Landy & Bergen 1991; Westland & Foster 1995).

(*f*) *Plan of experiments*

Psychophysical methods were used to measure human performance in both segmenting line-element textures and detecting line-element targets in stimuli containing large orientation heterogeneities. Two of the experiments replicated previous orientation-contrast experiments (Nothdurft 1991, 1993) in which individual line-elements in the stimuli were subjected to progressive variations in orientation; the other experiments used random variations of line-element orientation to introduce orientation heterogeneities. These experiments provided the data against which the performance of each of the models discussed in the next section was tested.

2. MODELS

(*a*) *Structure*

Figure 2 shows the general structure of the models. Each circle in the figure represents a module comprising a complete three-layer perceptron network (including the input layer). The modular structure reduced the complexity and training time of the ANNS: only one module of a given type needed to be trained, since the rest could be obtained by copying.

The models were designed to deal with line-element texture images like those of figure 1. Processing took place in three stages: stage 1 estimated the orientations of individual line-elements in the input image; stage 2 estimated the degree of orientation contrast at each location in the input image; and stage 3 classified the target region in the image as either horizontal or vertical, or, if the task was target detection, as either left or right (see figure 1). The individual processing units of stage 1 and stage 2 were intended to simulate some of the basic properties of cells in area V1 of visual cortex (see §1*d*). Stage 3 converted the output of the model into a form that could be compared with data from human observers, but it was not intended to model the cognitive processes underlying human decision processes. (A fuller account of the design, training, and testing of the models can be found in Schofield 1993.)

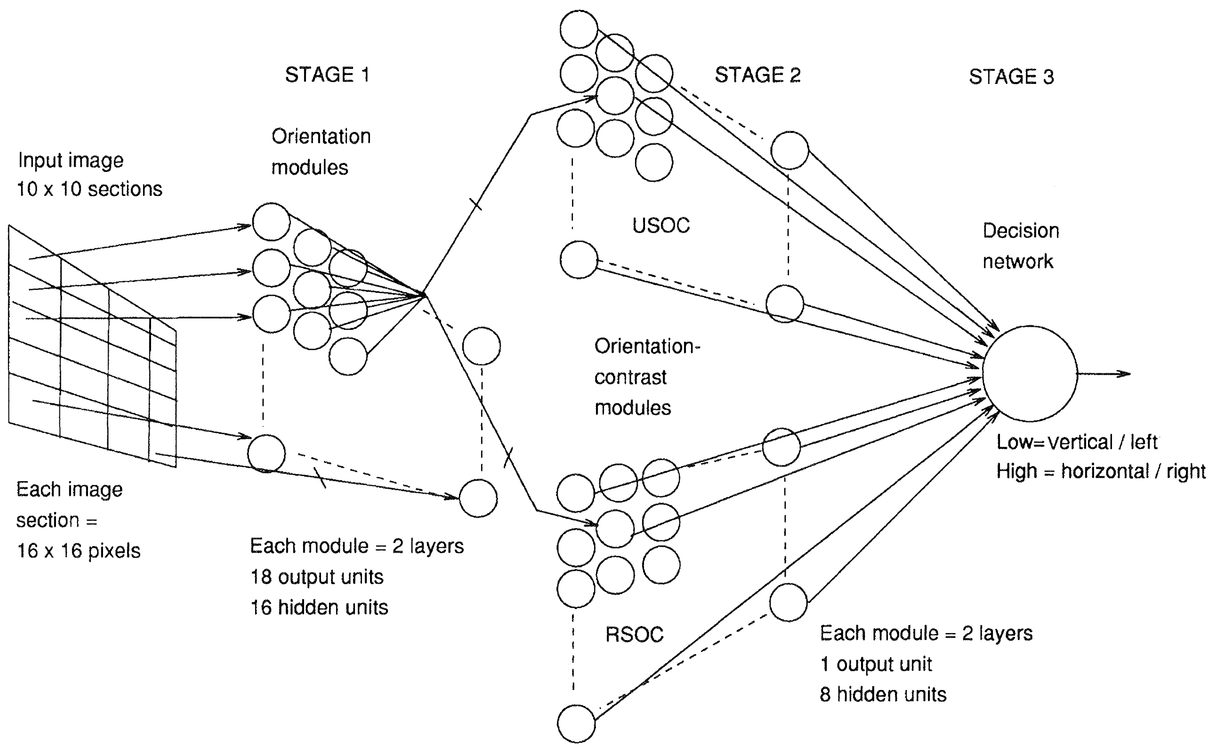

Figure 2. General structure of the models. USOC: uniform-suppressed orientation contrast; RSOC; random-suppressed orientation contrast.

(*b*) *Stage 1: orientation-sensitive modules*

The input image comprised 100 image sections of 16×16 pixels, each section containing a single 2×6 pixel line-element. Line elements were drawn with an anti-aliasing technique to reduce the effects of pixellation on sloping lines. The orientation of each line-element was estimated by one of 100 identical orientation-sensitive modules, each comprising 256 inputs (one for each pixel), 16 hidden units with sigmoidal output functions, and 18 output units with linear output functions. One such module was trained on images containing single line-elements and the other modules were generated by copying.

The module was trained on example lines with 12 orientations (15° intervals) positioned randomly within the input array. Stimuli were presented in random order. Output units were designed to be insensitive to the particular position of line-elements within their receptive fields. They were trained to have a variety of preferred orientations and orientation bandwidths, like standard-complex cells. Thus, 12 units were trained to have 15° bandwidths and preferred orientations at 15° intervals, four units to have 45° bandwidths and preferred orientations at 45° intervals, and two units to have 90° bandwidths and preferred orientations at 90° intervals: in all, a representative sample of complex-cell orientation-tuning properties (see Schiller *et al.* 1976). The response curves of the output units were bell-shaped. The network was trained to convergence.

The number of hidden units was determined by trial-and-error; the number eventually obtained (i.e. 16) was about the minimum necessary for the network as a whole to operate successfully. As the orientation-sensitive modules were trained, the hidden units, without being explicitly trained themselves, developed characteristics similar to those of simple cells in that they became tuned for both orientation and position (see Schofield 1993).

(*c*) *Stage 2: orientation-contrast-sensitive modules*

There were two types of orientation-contrast-sensitive modules (represented by the parallel pathways through stage 2 of figure 2). Each module had a single output unit, eight hidden units, all with sigmoidal output functions, and 162 inputs. The uniform-suppressed orientation-contrast-sensitive modules (USOC in figure 2) were trained to simulate the basic properties of uniform-suppressed orientation-contrast cells, and the random-suppressed orientation-contrast-sensitive modules (RSOC in figure 2) were trained to simulate the basic properties of random-suppressed orientation-contrast cells (see §1 *d*). Altogether there were 100 identical uniform-suppressed orientation-contrast-sensitive modules and 100 identical random-suppressed orientation-contrast-sensitive modules. The receptive fields of these modules overlapped. One module of each type was trained and the others were generated by copying.

Training was as follows. Nine line-elements arranged in a 3×3 array were pre-processed by nine copies of the orientation-sensitive module. Line-element orientations were chosen from the range 0°–180° at 5° intervals. The outputs of these modules were then passed to one of the two types of orientation-contrast-sensitive modules. The uniform-suppressed orientation-contrast-sensitive module was trained to give a high

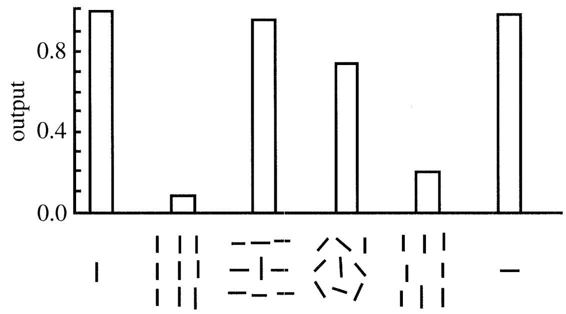

Figure 3. Response of the uniform-suppressed orientation-contrast-sensitive module to various line-element arrays. Image types are represented by the icons under each bar. The height of each bar represents the output-unit's average response to images of the type represented by the icon.

response to single line-elements at any orientation and to line-element arrays in which a single line-element was surrounded by line-elements orthogonal to it. It was trained to give a low response to arrays of uniformly oriented line-elements, to a blank field, and to arrays containing no central line-element. It was also trained to give a low response to arrays in which line-element orientations were chosen randomly from a restricted range ($\pm 20°$). It was not given any explicit training for arrays in which orientations were chosen randomly from the full range, although after training it was found to give a moderately high response to such images. Stimuli were presented in random order, with all stimulus types equally represented. The network was trained to convergence. Figure 3 shows the responses of this output unit to different types of images represented by the icons below each column.

The random-suppressed orientation-contrast-sensitive module was trained to give a high response to single line-elements, to line-element arrays in which a single line-element was surrounded by line-elements orthogonal to it, and to arrays of line-elements with a small amount of random orientation (see later). It was trained to give a low response to arrays with completely random line-element orientations, to a blank field, and to arrays containing no central line-element. It was not explicitly trained to give a high response to arrays of uniformly oriented line-elements, although after training it was found to do so. Stimuli were presented in random order, with all stimulus types equally represented. The network was trained to convergence. The response of this module to the different types of images is shown in figure 4.

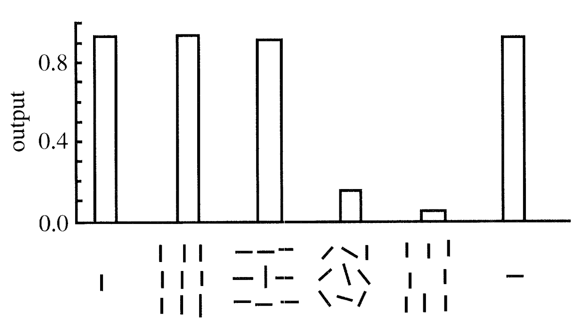

Figure 4. Response of the random-suppressed orientation-contrast-sensitive module to various line-element arrays. Image types are represented by the icons under each bar. The height of each bar represents the output-unit's average response to images of the type represented by the icon.

The responses of the modules (see figures 3 and 4) are more categorical than the average responses of the cell types they modelled (Knierim & Van Essen 1992); but they did capture the gross functionality of their biological counterparts. Notice that the responses shown in figures 3 and 4 refer to the output units of the orientation-contrast-sensitive modules only. The responses of units in the hidden layers were difficult to characterize, but they showed a range of orientation-contrast sensitivities and many were sensitive to absolute orientation.

(*d*) *Model variants*

To test the contributions of the two types of orientation-contrast-sensitive module to overall performance, the model was assembled in four variants. Model 1 had no orientation-contrast-sensitive modules at all, the outputs of the orientation-sensitive modules being fed directly to stage 3 (the decision network). Model 2 had orientation-sensitive modules and uniform-suppressed orientation-contrast-sensitive modules only. Model 3 had orientation-sensitive modules and random-suppressed orientation-contrast-sensitive modules only. Model 4 had orientation-sensitive modules and both types of orientation-contrast-sensitive modules. There were no direct connections between stages 1 and 3 in models 2, 3, and 4.

(*e*) *Stage 3: decision logic*

The decision network comprised two adaptive units in a single layer. The final output was set high or low depending on which of the two units had the greater response. A unique version of the network was developed for each model. The network was trained to classify the target regions as either horizontal or vertical, and the target line-element as either left or right, depending on the task. Thus one decision unit was trained to give a high response to a horizontal (or left) stimulus, the other a high response to a vertical (or right) stimulus.

The stage 3 variants were trained on the outputs of stage 2 of the corresponding model (because model 1 did not have any stage 2 modules, its decision network was trained on the outputs of stage 1). Models 1, 2, and 3 were used only in the texture-segmentation experiments and so it was necessary to develop only one version of the decision network for each of the models. Two versions of the final stage were produced for model 4: one version was trained to segment texture images like those of figure 1*a*; the other was trained to detect target elements in images like those of figure 1*e*. During training, the models were presented with homogeneous types of images (each containing only two line-element orientations) that they were expected to process (like those in figures 1 *a*, *e*), with individual line-elements chosen from the range given in (*c*). Each variant of stage 3 was trained to make the best use of the information provided by the earlier stages. None was trained explicitly to deal with heterogeneous types of images (each containing more than two line-element orientations) because this could have allowed the decision stage to circumvent the deficiencies in stages 1 and 2 rather than reflecting their performances accurately.

(*f*) *Testing*

The models were implemented as software simulations on standard single-processor computers (Sun Microsystems Inc., California, U.S.A.; SPARCstation family). The images presented to the models were, apart from being spatially quantized, the same as those presented to the human observers. The outputs of the models were recorded and then analysed also in the same way as for the human observers. Models 1, 2, and 3 were not used in the target-detection experiments (experiments 5, 6, and 7) because they were found to be incapable of modelling human performance in the texture-segmentation experiments (experiments 1, 2, and 3). These models were also excluded from one of the texture-segmentation experiments (experiment 4) for the same reason. Model 4 was used in all experiments.

3. PSYCHOPHYSICAL METHODS

There were four texture-segmentation experiments and three target-detection experiments. The same procedure was used in each. Observers were presented with displays similar to those shown in figure 1 and were required to report either the orientation of the central texture region (horizontal or vertical) or the location of the target line-element (left or right).

Stimuli were drawn on the screen of an X-Y display oscilloscope (Hewlett-Packard, U.S.A.; Type 1321A, P4 sulphide phosphor) driven by a vector-graphics generator (Sigma Electronic Systems, U.K.; QVEC 2150) and additional digital-to-analogue converters, in turn controlled by a laboratory computer (see, for example, Foster & Ward 1991*a* for details). The angular subtense of each line-element was about 0.1°, and of the whole line-element array 5.0° by 5.0°. The display was viewed binocularly at 170 cm through a view-tunnel, which produced a uniformly illuminated white background of luminance 40 cd m^{-2} , on which the stimuli appeared superimposed. At the beginning of each experimental session, the observer, using a neutral density filter, set the stimulus luminance to tenfold luminance-increment threshold.

The observer fixated a centrally located cross on the screen and initiated the trial by pressing a button switch connected to the computer. The cross disappeared and the test stimulus then appeared for 100 ms, after which the observer responded using one of two button switches connected to the computer. Observers were encouraged to respond as quickly as was consistent with accuracy. There were 49 different stimulus conditions in each of the seven experiments and observers contributed at least 40 responses in each condition. The experiments were conducted in runs of 245 trials in which the ordering of the trials was chosen at random, subject to the constraint that each condition was tested five times in total.

There were five observers: one was the first author, but the others were unaware of the aims of the experiment and were paid for their time. They were male, aged 21–27 years, with normal or corrected-to-normal visual acuity, and negligible astigmatism. At least two observers participated in each experiment, but the first author was the only one to participate in all of the experiments.

4. EXPERIMENT 1: SEGMENTATION WITH PROGRESSIVE ORIENTATION VARIATIONS

(*a*) *Method*

The stimuli were similar to those used in previous orientation-contrast studies (Nothdurft 1991). Displays consisted of 100 line-elements arranged in a 10×10 array. For each display a rectangular target region (6×2 line-elements) was defined in the centre of the display, as illustrated in figure 1 *b*, and oriented either horizontally or vertically. For each display, the 'within-region' orientation increment $\Delta\theta_{\mathrm{W}}$ ranged over 0°–60° at 10° intervals. With the top left line-element as origin, the orientation of each line-element was determined by adding $\Delta\theta_{\mathrm{W}}$ to the orientation of the adjacent line-element on the left or in the row above. The orientations of the line-elements within the target region were subjected to an additional increment $\Delta\theta_{\mathrm{T}}$, the value of which also ranged over 0°–60° at 10° intervals. The difference $\Delta\theta_{\mathrm{TB}}$ in orientation between a line-element in the target region and any adjacent line-element in the background was therefore given by:

$$\Delta\theta_{\mathrm{TB}} = \Delta\theta_{\mathrm{W}} + \Delta\theta_{\mathrm{T}}.$$

The polarity of the orientation increment was reversed for line-elements below or to the right of the central target region so that $\Delta\theta_{\mathrm{TB}}$ remained constant on all sides of the target region.

For each value of $\Delta\theta_{\mathrm{W}}$ the percent correct score for each value of $\Delta\theta_{\mathrm{TB}}$ was recorded, thereby defining a psychometric function. Each of the seven psychometric functions (one for each value of $\Delta\theta_{\mathrm{W}}$) was transformed by the inverse of a cumulative Gaussian (probit) function and a quadratic function fitted to the transformed data points (a polynomial of order at least two was needed because orientation is a circular variable; the polynomial was not replaced by a circular function, however, because the periodicity of the underlying detection mechanisms was unknown). The threshold value of $\Delta\theta_{\mathrm{TB}}$ for a criterion score of 76% correct, corresponding in signal-detection theory to a discrimination index value d' of 1.0, was then determined from each of the fitted curves. Standard deviations of these threshold values were estimated by a bootstrap method (Foster & Bischof 1991).

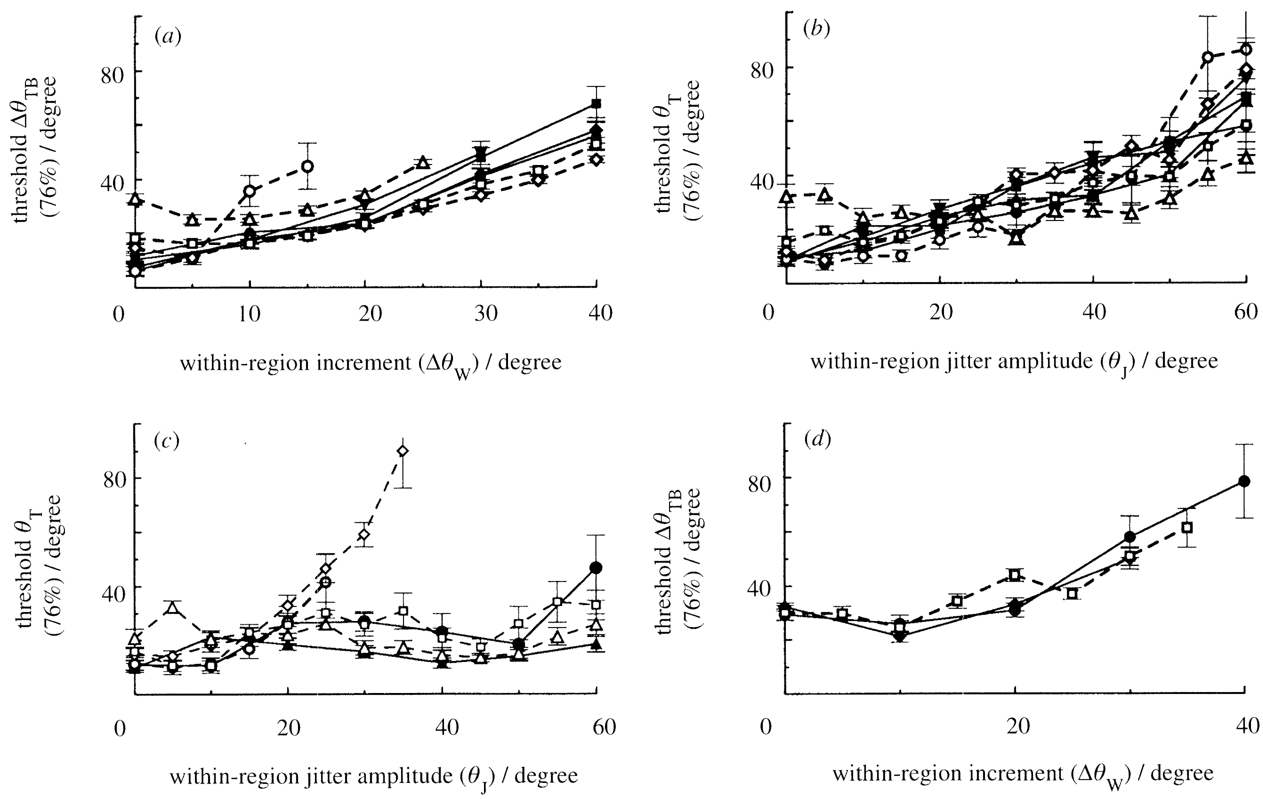

Figure 5. Results for human observers (continuous lines) and models (broken lines) in texture-segmentation experiments: (*a*) experiment 1, progressive orientation variations; (*b*) experiment 2, random orientation variations; (*c*) experiment 3, grouped orientation variations; (*d*) experiment 4, progressive orientation variations with jitter. Threshold difference in orientation between target and background required for a 76% correct score is plotted against the amplitude of within-region orientation variation. ○– – –○ model 1; ◇– – –◇ model 2; △– – –△ model 3; □– – –□ model 4. ◆——◆ R.P.M.; ●——● A.J.S.; ■——■ D.E.; ▲——▲ G.S.; ▼——▼ P.T.

(*b*) *Results and comment*

Figure 5*a* shows the results. Data from human observers are shown by continuous lines. Despite the large overlap in the orientations of the line-elements in each of the two regions, observers were able to segment displays in the presence of large orientation heterogeneities. As the within-region orientation increment $\Delta\theta_W$ increased, threshold values of the difference $\Delta\theta_{TB}$ between target and background orientation increased. These results are in broad agreement with previous findings (Nothdurft 1991).

Data from the models are shown by the broken lines. Threshold values of $\Delta\theta_{TB}$ from model 1 (open circles) and from model 3 (open triangles) were higher than those from human observers, but threshold values from both model 2 (open diamonds) and model 4 (open squares) followed those for human observers reasonably well.

5. EXPERIMENT 2: SEGMENTATION WITH RANDOM ORIENTATION VARIATIONS

(*a*) *Method*

Because of the way in which line-element orientation was varied progressively over successive rows of the line-element arrays in experiment 1, the visual appearance of these stimuli was that of a 'flow pattern'. It was possible that the disruptions in these patterns by the target region provided the principal cue to discrimination performance. To eliminate flow patterns, line-element orientation was randomized with values drawn from a uniform distribution. The extent of the randomization in each array was defined by a jitter amplitude θ_J that ranged over 0°–60° at 10° intervals. The orientations of the background line-elements were chosen randomly from the range $-\theta_J$ to $+\theta_J$ at 5° intervals, and the orientations of the target line-elements randomly from the range $\theta_T-\theta_J$ to $\theta_T+\theta_J$ at 5° intervals. The difference between the mean orientations of the two regions was θ_T, and this value, like θ_J, ranged over 0°–60° at 10° intervals. The threshold value of θ_T for a score of 76% correct was determined, as before, for each amplitude θ_J of within-region heterogeneity. Figure 1*c* shows an example display.

(*b*) *Results and comment*

Figure 5*b* shows the results. Data from human observers are shown by continuous lines. Threshold values of the difference $\Delta\theta_T$ between mean target and mean background orientations increased with increasing jitter amplitude θ_J, but, as in experiment 1, segmentation was possible even when there was large overlap in the orientations of the line-elements in the two regions.

Data from the models are shown by the broken lines. Threshold values of $\Delta\theta_T$ from model 1 (open circles) were too high at large jitter amplitudes; threshold values from model 3 (open triangles) showed the opposite effect and were also too high at small jitter amplitudes. Threshold values from models 2 (open diamonds) and 4 (open squares) followed those for human observers well.

6. EXPERIMENT 3: SEGMENTATION WITH GROUPED ORIENTATION VARIATIONS

(*a*) *Method*

According to one general theory of the visual processing of multi-element displays (Duncan & Humphreys 1989), line-element images with few orientations should be easier to segment than those with many orientations, even if the ranges of orientations present are the same. The suggestion is based on the hypothesis that 'similar' elements can be grouped perceptually and can therefore be processed more efficiently than elements that cannot be grouped (care has to be taken in defining similarity, as some apparently similar elements group less well than some apparently less similar elements: see, for example, Beck 1966, 1974; Treisman 1991). This suggestion was explored within the present framework using displays of the kind shown in figure 1*d*. As in experiment 2, the jitter amplitude θ_J and mean target orientation θ_T ranged over 0°–60° at 10° intervals. The orientations of individual line-elements were then chosen randomly from the extremes of the possible orientation ranges: background line-elements could have one of two orientations $-\theta_J$ or $+\theta_J$, and line-elements in the target region either $\theta_T-\theta_J$ or $\theta_T+\theta_J$. In this way, the span of orientations for any stimulus combination was the same as for the equivalent combination in experiment 2, but no image contained more than four orientations.

(*b*) *Results and comment*

Figure 5*c* shows the results. Data from human observers are shown by continuous lines and from the models by broken lines. Threshold values of the difference θ_T between mean target and mean background orientations increased only slightly with increasing jitter amplitude θ_W and they were markedly lower than in experiment 2, suggesting that observers were able to exploit grouping effects. (The role of grouping effects in this experiment has been addressed elsewhere in an ideal-observer analysis; see Schofield 1993.)

Threshold values of θ_T from models 1 (open circles) and 2 (open diamonds) increased more rapidly with jitter amplitude θ_W, but threshold values from models 3 (open triangles) and 4 (open squares) both followed those for human observers well, although those for model 4 were slightly better (especially at small jitter amplitudes).

From figures 5*a*–*c*, it is evident that model 4 was the only one that accounted well for human performance in all three experiments. With a combination of the two types of orientation-contrast unit, it produced variations in threshold values similar to those for human observers with line-element images containing progressive, random, progressive and random, and grouped orientation variations.

7. EXPERIMENT 4: SEGMENTATION WITH PROGRESSIVE ORIENTATION VARIATIONS AND ADDITIONAL ORIENTATION JITTER

(*a*) *Method*

A different technique was used here to eliminate the flow patterns in the arrays used in experiment 1: the displays were the same but an additional orientation jitter was added to the orientation of each line-element. This jitter amplitude was chosen randomly from the range $-30°$ to $+30°$ and was applied to all images regardless of the underlying within-region increment $\Delta\theta_W$. As in experiment 1, threshold values of $\Delta\theta_{TB}$ were determined as a function of $\Delta\theta_W$. Only model 4 was tested in this experiment because, as was noted earlier, it was the only one to capture successfully human performance in experiments 1, 2 and 3.

(*b*) *Results and comment*

Figure 5*d* shows the results. Data from human observers are shown by continuous lines and from the model by broken lines. Threshold values of the differences $\Delta\theta_{TB}$ between mean target and mean background orientations were similar for the two and markedly higher than in experiments 1 and 2, presumably a consequence of the increased orientation heterogeneity.

It may be noted that there are local non-smooth variations in the performance of this and other models, especially with low jitter amplitudes and orientation increments. A contributory factor may have been the varying densities at which orientation was sampled: 15° in training but 5° in testing.

8. EXPERIMENTS 5, 6 AND 7: TARGET DETECTION

The next sections describe three target-detection experiments corresponding to experiments 1, 2, and 3. Because of the inadequate performance of models 1–3 in the texture-segmentation experiments, only model 4 was tested in the target-detection tasks.

(*a*) *Method*

In each of these experiments the central texture region of the previous experiments was replaced by a single target line-element, which could be placed at one of two locations each side of the vertical midline of the images. The stimuli in experiment 5 were constructed in the same way as in experiment 1; that is, the orientation of each line-element was determined by adding a fixed increment $\Delta\theta_W$ to the orientation of the adjacent line-element. Unlike the texture-

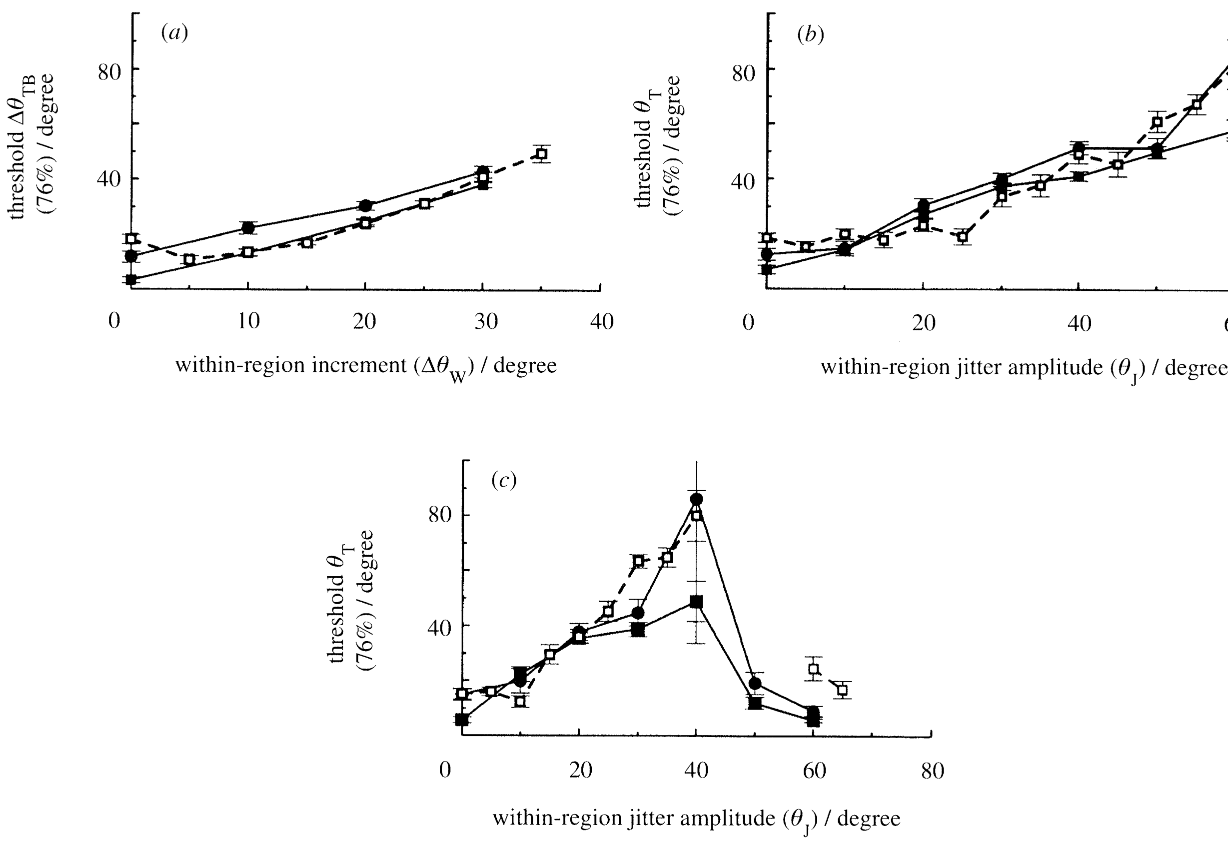

Figure 6. Results for human observers (continuous lines) and model 4 (broken lines) in target-detection experiments: (*a*) experiment 5, progressive orientation variations; (*b*) experiment 6, random orientation variations; (*c*) experiment 7, grouped orientation variations. Threshold difference in orientation between target and background required for a 76 % correct score is plotted against the amplitude of within-region orientation variation. ●——● A.J.S.; ■——■ J.A.; □– – –□ model 4.

segmentation experiment, however, the polarity of the orientation increment was changed in the region of both target locations, so that the standard polarity change (see experiment 1, §4*a*) could not be used to identify the target location. The stimuli used in this experiment were similar to those used in previous target-detection experiments (Nothdurft 1991, 1993). Figure 1*f* shows an example display. As in experiment 1, threshold values of the difference $\Delta\theta_{TB}$ between the orientation of the line-element target and the orientation of any adjacent line-element in the background were determined as a function of $\Delta\theta_W$.

The stimuli of experiment 6 were constructed in the same way as in experiment 2, but no jitter was added to the target line-element orientation θ_T. Figure 1*g* shows an example display.

The stimuli in experiment 7 were constructed in a similar way to those in experiment 3. Background line-element orientations were chosen randomly from the two orientations $-\theta_J$ and $+\theta_J$, but no orientation jitter was added to the target line-element orientation θ_T. Figure 1*h* shows an example display. Because of this design, an unplanned cue was generated in displays with large θ_J and low θ_T. In those displays, background line-element orientations were far from the mean background orientation 0° and the target line-element orientation θ_T was close to the mean background orientation; target line-element orientation was therefore unique and clearly different from each of the background orientations.

(*b*) *Results and comment*

Figure 6 shows the results of the three target-detection experiments. Data from human observers are shown by the continuous lines and from model 4 by broken lines. Threshold values of the difference between target and background orientations increased with each of the independent orientation variables $\Delta\theta_W$ and θ_J, but observers could still detect the targets even in the presence of large variations in line-element orientation. Observers did, however, find the tasks more difficult than texture segmentation: threshold values with large $\Delta\theta_W$ and θ_J were in general higher than the equivalent threshold values in figure 5, most notably for grouped orientation variations (experiments 3 and 7), indicating the ineffectiveness of grouping strategies in this kind of target-detection task (cf. Alkhateeb *et al.* 1990; Wolfe & Friedman-Hill 1992). The unintentional cue mentioned earlier with large jitter amplitudes θ_J in experiment 7 presumably accounted for the very low threshold values when $\theta_J >$ 40° (see figure 6*c*) obtained for both human observers and the model. Although the model failed to reach threshold for jitter amplitudes θ_J in the region of 50°, it still performed well at larger amplitudes.

9. DISCUSSION

Of the four types of analogue neural networks considered here, the most successful, namely model 4, was able to account for human performance in segmenting line-element regions and detecting line-element targets within line-element arrays containing a range of orientations. Several distinct types of ANN module appeared to be necessary for processing these images, a result which, if relevant to the structure of biological vision systems, suggests that the existence of various types of cortical cells, with different line-element-orientation and orientation-contrast characteristics, is an important factor in determining human performance in these tasks. Although it might have been possible to produce a similar performance with hard-wired models, the choice of parameters for the orientation-contrast units would not have been straightforward. The use of ANNs avoided the need to define explicitly the connection strengths for these units.

Model 4 contained one type of orientation-sensitive module and two types of orientation-contrast-sensitive modules: those that suppressed responses to arrays of line-elements of uniform orientation and those that suppressed responses to arrays of line-elements of random orientation. By design, the outputs of those two types of orientation-contrast-sensitive modules were functionally similar to many of the orientation-contrast-sensitive cells found in area V1 responding to arrays of line-elements; and, without being explicitly trained, the hidden units within the orientation-sensitive modules preceding the orientation-contrast-sensitive modules acquired positional and orientational characteristics similar to those of simple cells.

Model 4 yielded thresholds for the difference between target and background orientations in texture segmentation and in line-target detection that closely matched those from human observers, both qualitatively and quantitatively. In particular, for texture segmentation with progressive and random orientation variations, thresholds increased with increasing orientation variance (see figure 5*a*, *b*), but with grouped orientation variations, thresholds increased more slowly (see figure 5*c*). For line-element detection with progressive and random orientation variations, thresholds increased with increasing orientation variance (see figure 6*a*, *b*), at about the same rate as for texture segmentation, but with grouped orientation variations, thresholds first increased more rapidly with increasing orientation variance than in texture segmentation, and then decreased more rapidly still (see figure 6*c*).

In principle, when line-elements can be grouped in a manner that is compatible with the requirements of the segmentation or detection task, performance should be better than when line-elements cannot be grouped. The present results suggest that this principle may hold for segmenting visual textures on the basis of line-element orientation, but not for detecting line-element targets in the equivalent backgrounds.

Why were two types of orientation-contrast-sensitive module necessary? It seems that line-element arrays with different types of orientation heterogeneity have to be processed differently to achieve high texture-segmentation and high target-detection performance: images with progressive variations in line-element orientation within regions are best processed by uniform-suppressed modules, whereas images in which orientations vary non-progressively but are similar within regions are best processed by random-suppressed modules.

The design of model 4 was partly motivated by the assumed division of cortical orientation-contrast cells into distinct classes, a division that is, to some extent, uncertain (Knierim & Van Essen 1992). This uncertainty might be resolved if neurophysiological data were available on the responses of observed cell types to grouped stimuli, such as those used in experiment 3. In containing two distinct types of orientation-contrast-sensitive module, model 4 differs from some hard-wired models of texture segmentation (e.g. Fogel & Sagi 1989; Malik & Perona 1990; Landy & Bergen 1991; Westland & Foster 1995) and from some partially adaptive models (e.g. Mesrobian & Skrzypek 1995), all of which have used orientation-sensitive units, or filters followed by orientation-contrast units, or orientation-gradient estimators.

Finally, it should be noted that although model 4 performed the given tasks well, its training with a specific, albeit large, class of line-element images may not have been optimal for other more general tasks, such as segmenting natural scenes. Moreover, as with other models of psychophysical performance, the success of model 4 does not imply its uniqueness: other neurophysiologically relevant ANNs might be constructed that achieved similar levels of performance in processing the kinds of line-element images considered here.

We thank L. M. Doherty, M. G. A. Thomson and S. Westland for critical reading of the manuscript. A.J.S. was supported by a training award from the U.K. MRC.

REFERENCES

Alkhateeb, W. F., Morris, R. J. & Ruddock, K. H. 1990 Effects of stimulus complexity on simple spatial discriminations. *Spatial Vision* **5**, 129–141.

Beck, J. 1966 Effect of orientation and of shape similarity on perceptual grouping. *Percept. Psychophys.* **1**, 300–302.

Beck, J. 1972 Similarity grouping and peripheral discriminability under uncertainty. *Am. J. Psychol.* **85**, 1–19.

Beck, J. 1973 Similarity grouping of curves. *Percept. motor skills* **36**, 1331–1341.

Beck, J. 1974 Relation between similarity grouping and peripheral discriminability. *J. exp. Psychol.* **102**, 1145–1147.

Beck, J. 1982 Textural segmentation. In *Organization and representation in perception* (ed. J. Beck), chapter 15. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Beck, J. 1983 Textural segmentation, second-order statistics, and textural elements. *Biol. Cyber.* **48**, 125–130.

Dean, A. F. & Tolhurst, D. J. 1983 On the distinctness of simple and complex cells in the visual cortex of the cat. *J. Physiol.* **344**, 305–325.

De Valois, R. L., Yund, E. W. & Hepler, N. 1982 The orientation and direction selectivity of cells in macaque visual cortex. *Vis. Res.* **22**, 531–544.

Dreher, B. 1972 Hypercomplex cells in the cat's striate cortex. *Invest. Ophthalmol.* **11**, 355–356.

Duncan, J. & Humphreys, G. W. 1989 Visual search and stimulus similarity. *Psychol. Rev.* **96**, 433–458.

Fogel, I. & Sagi, D. 1989 Gabor filters as texture discriminator. *Biol. Cyber.* **61**, 103–113.

Foster, D. H. & Mason, R. J. 1980 Irrelevance of local position information in visual adaptation to random arrays of small geometric elements. *Perception* **9**, 217–221.

Foster, D. H. & Bischof, W. F. 1991 Thresholds from psychometric functions: Superiority of bootstrap to incremental and probit variance estimators. *Psychol. Bull.* **109**, 152–159.

Foster, D. H. & Ward, P. A. 1991*a* Asymmetries in oriented-line detection indicate two orthogonal filters in early vision. *Proc. R. Soc. Lond.* B **243**, 75–81.

Foster, D. H. & Ward, P. A. 1991*b* Horizontal-vertical filters in early vision predict anomalous line-orientation identification frequencies. *Proc. R. Soc. Lond.* B **243**, 83–86.

Foster, D. H. & Westland, S. 1995 Orientation contrast vs orientation in line-target detection. *Vis. Res.* **35**, 733–738.

Gilbert, C. D. 1977 Laminar differences in receptive field properties of cells in cat primary visual cortex. *J. Physiol.* **268**, 391–421.

Gurnsey, R. & Browse, R. A. 1987 Micropattern properties and presentation conditions influencing visual texture discrimination. *Percept. Psychophys.* **41**, 239–252.

Hubel, D. H. & Wiesel, T. N. 1968 Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* **195**, 215–243.

Javadnia, A. & Ruddock, K. H. 1988 The limits of parallel processing in the visual discrimination of orientation and magnification. *Spatial Vision* **3**, 97–114.

Julesz, B. 1962 Visual pattern discrimination. *IRE Trans. Inform. Theor.* **IT-8**, 84–92.

Julesz, B. 1980 Spatial nonlinearities in the instantaneous perception of textures with identical power spectra. *Phil. Trans. R. Soc. Lond.* B **290**, 83–94.

Julesz, B. 1981*a* A theory of preattentive texture discrimination based on first-order statistics of textons. *Biol. Cyber.* **41**, 131–138.

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

Julesz, B. 1984 A brief outline of the texton theory of human vision. *Trends NeuroSci.*, 41–45.

Knierim, J. J. 1991 Neural responses to texture patterns in area V1 of the alert monkey. PhD thesis: California Institute of Technology.

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

Landy, M. S. & Bergen, J. R. 1991 Texture segregation and orientation gradient. *Vis. Res.* **31**, 679–691.

Lehky, S. R. & Sejnowski, T. J. 1988 Network model of shape-from-shading: neural function arises from both receptive and projective fields. *Nature, Lond.* **333**, 452–454.

Lehky, S. R. & Sejnowski, T. J. 1990 Neural network model of visual cortex for determining surface curvature from images of shaded surfaces. *Proc. R. Soc. Lond.* B **240**, 251–278.

Malik, J. & Perona, P. 1990 Preattentive texture discrimination with early vision mechanisms. *J. opt Soc. Am.* A **7**, 923–932.

Marendaz, C., Stivalet, P., Barraclough, L. & Walkowiac, P. 1993 Effect of gravitational cues on visual search for orientation. *J. exp. Psychol. Hum. Percept. Perform.* **19**, 1266–1277.

Marendaz, C., Stivalet, P. & Genon, D. 1991 Coding of orientation in early vision: search asymmetry revisited. *Eur. Bull. cog. Psychol.* **11**, 427–440.

Meigen, T., Lagrèze, W.-D. & Bach, M. 1994 Asymmetries in preattentive line detection. *Vis. Res.* **34**, 3103–3109.

Mesrobian, E. & Skrzypek, J. 1995 Segmenting textures using cells with adaptive receptive fields. *Spatial Vision* **9**, 163–190.

Moorhead, I. R., Haig, N. D. & Clement, R. A. 1989 An investigation of trained neural networks from a neurophysiological perspective. *Perception* **18**, 793–803.

Nakayama, K., Silverman, G. H., MacLeod, D. I. A. & Mulligan, J. 1985 Sensitivity to shearing and compressive motion in random dots. *Perception* **14**, 225–238.

Neisser, U. 1967 *Cognitive psychology*. Englewood Cliffs, New Jersey: Prentice-Hall.

Nothdurft, H. C. 1985*a* Orientation sensitivity and texture segmentation in patterns with different line orientation. *Vis. Res.* **25**, 551–560.

Nothdurft, H. C. 1985*b* Sensitivity for structure gradient in texture discrimination tasks. *Vis. Res.* **25**, 1957–1968.

Nothdurft, H. C. 1990 Texton segregation by associated differences in global and local luminance distribution. *Proc. R. Soc. Lond.* B **239**, 295–320.

Nothdurft, H. C. 1991 Texture segmentation and pop-out from orientation contrast. *Vis. Res.* **31**, 1073–1078.

Nothdurft, H.-C. 1993 The role of features in preattentive vision: comparison of orientation, motion and color cues. *Vis. Res.* **33**, 1937–1958.

Olson, R. K. & Attneave, F. 1970 What variables produce similarity grouping? *Am. J. Psychol.* **83**, 1–21.

Rubenstein, B. S. & Sagi, D. 1990 Spatial variability as a limiting factor in texture-discrimination tasks: implications for performance asymmetries. *J. opt. Soc. Am.* A **7**, 1632–1643.

Rumelhart, D. E., McClelland, J. L. & The PDP Research Group 1986 *Parallel distributed processing*, vol. 1: foundations. Cambridge, Massachusetts: MIT Press.

Schiller, P. H., Finlay, B. L. & Volman, S. F. 1976 Quantitative studies of single-cell properties in monkey striate cortex. II. Orientation specificity and ocular dominance. *J. Neurophysiol.* **39**, 1320–1333.

Schofield, A. J. 1993 Neural network models for texture segmentation and target detection. PhD thesis: Keele University.

Schofield, A. J. & Foster, D. H. 1993 A modular network model for texture segmentation based on orientation contrast. In *Neural computing research and applications* (ed. G. Orchard), pp. 53–61. Bristol: Institute of Physics Publishing.

Treisman, A. 1977 Focused attention in the perception and retrieval of multidimensional stimuli. *Percept. Psychophys.* **22**, 1–11.

Treisman, A. 1985 Preattentive processing in vision. *Computer Vis. Graphics Image Process.* **31**, 156–177.

Treisman, A. 1991 Search, similarity, and integration of features between and within dimensions. *J. exp. Psychol. Hum. Percept. Perform.* **17**, 652–676.

Treisman, A. & Gormican, S. 1988 Feature analysis in early vision: Evidence from search asymmetries. *Psychol. Rev.* **95**, 15–48.

Treisman, A. & Souther, J. 1985 Search asymmetry: A diagnostic for preattentive processing of separable features. *J. exp. Psychol. Gen.* **114**, 285–310.

Van Essen, D. C., DeYoe, E. A., Olavarria, J. F., Knierim, J. J., Fox, J. M., Sagi, D. & Julesz, B. 1989 Neural responses to static and moving texture patterns in visual cortex of the macaque monkey. In *Neural mechanisms of visual perception* (ed. D. M.-K. Lam, D. Gilbert & C. D. Gilbert), pp. 137–154. The Woodlands, Texas: Portfolio Publishing Company.

Westland, S. & Foster, D. H. 1995 Optimized model of oriented-line-target detection using vertical and horizontal filters. *J. opt. Soc. Am.* A **12**, 1617–1622.

Wolfe, J. M. 1992 "Effortless" texture segmentation and "parallel" visual search are *not* the same thing. *Vis. Res.* **32**, 757–763.

Wolfe, J. M. & Friedman-Hill, S. R. 1992 On the role of symmetry in visual search. *Psychol. Sci.* **3**, 194–198.

Received 1 December 1994; accepted 15 June 1995