Optimized model of oriented-line-target detection using vertical and horizontal filters

Stephen Westland and David H. Foster

Department of Communication and Neuroscience, Keele University, Staffordshire ST5 5BG, UK

A line-element target differing sufficiently in orientation from a background of line elements can be visually detected easily and quickly; orientation thresholds for such detection are lowest when the background elements are all vertical or all horizontal. A simple quantitative model of this performance was constructed from three processing stages: (1) linear filtering by two classes of anisotropic filters, (2) nonlinear point transformation, and (3) estimation of a signal-to-noise ratio based on responses to images with and without a target. A Monte Carlo optimization procedure (simulated annealing) was used to determine the model parameter values required for providing an accurate description of psychophysical data on orientation increment thresholds.

1. INTRODUCTION

There are many possible cues to the visual detection of a target in a complex background, but in monochromatic, static, monocularly viewed images one of the most important is the orientation of lines and edges: providing that the difference in orientation of target and background lines is sufficiently large, detection can be achieved speedily and without the need for thought or attention.[1–10] The underlying visual mechanisms are assumed to operate in an essentially parallel way over the visual field, as part of an early (preattentive or distributed-attention) level of visual processing.[2–5,11,12]

There have been few systematic psychophysical measurements of the orientation dependence of this performance. One parametric investigation[7] of the detectability of line-element targets in random arrays of line elements (like those illustrated in Fig. 1) has shown that performance is anisotropic: the angular difference between target and background line elements at threshold varies markedly with background orientation. A first-order theoretical analysis[7] of those data led to the proposal that the early visual processes determining performance in this particular task are dominated by two classes of orientation-sensitive filters, whose major axes are near the vertical and the horizontal with respect to the frontoparallel plane of view.[7,10]

Based on that analysis, this study sets forth a simple, quantitative model for line-element-detection performance. The model comprised three stages: (1) linear filtering by two classes of anisotropic filters, (2) nonlinear point transformation, and (3) estimation of a signal-to-noise ratio based on responses to images with and without a target. An optimization procedure—simulated annealing—was used to determine the values of the parameters of the model so that it provided an accurate description of data on orientation increment thresholds. A report of some preliminary results from this study has been published previously.[13]

2. PSYCHOPHYSICAL DATA AND ANALYSIS

The psychophysical data to be fitted were taken from a previous investigation.[7] Relevant details are, briefly, as follows. Observers viewed achromatic stimulus images consisting of a random array of 20 line elements generated on a vector-graphics $X-Y$ display. In half of the images 19 of the line elements—forming the background—had the same orientation (angle θ with respect to the vertical), and one line element—the target—had a different orientation (angle $\theta + \Delta\theta$), as shown in Fig. 1; both θ and $\Delta\theta$ varied randomly over trials. (In the experiment the line elements appeared white against a gray background.) In the other half of the images the line elements all had the same orientation (angle θ), which also varied randomly over trials. The images were each presented briefly, and the observer had to decide whether a target was present. For a given level of detection performance, corresponding to a fixed criterion level defined by the discrimination index d' from signal-detection theory,[14] increment threshold $\Delta\theta$ was determined as a function of θ. Data were pooled over a group of ten observers. A representative portion of the data is plotted in Fig. 2 (open circles; $d' = 0.2$). Increment thresholds were smallest when the background elements were oriented along the vertical and the horizontal. Notice that $\Delta\theta$ and θ were measured in the same direction, and therefore asymmetries in the dependence of $\Delta\theta$ on θ are not unexpected.[3,7,9,10] (The data represent a moderately coarse sampling of the increment-threshold function for orientation, and, despite the large number of observations, the standard deviations associated with each threshold are sufficiently large that the pattern of asymmetry is not precisely defined.)

Previous analysis[7] of these data yielded estimates of the orientation-tuning curves of the proposed underlying filters: their axes were oriented at $-4 \pm 8°$ and $83 \pm 6°$ with respect to the vertical (positive anticlockwise), and their half-height half-widths were approximately $30°$.

Fig. 1. Typical stimulus image. The orientation of the background line elements is 15°; of the target line element, 45°.
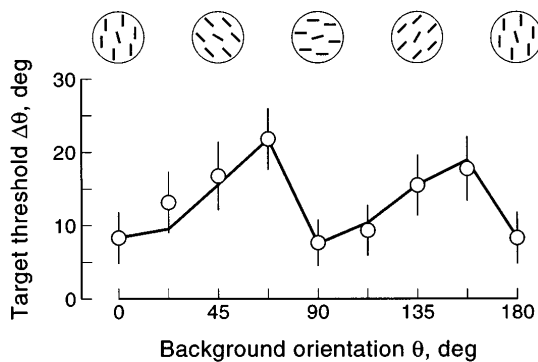

Fig. 2. Increment threshold for the detection of a target line element within a field of background line elements as a function of orientation of the background elements. Open circles: psychophysical data,[7] means ±1 standard error of the mean; the point at 180° is identical to the point at 0°. Solid curve: data from the optimized model.

These tuning curves very closely matched—in shape, position, and relative heights—the orientation-tuning curves of certain vertical and horizontal components obtained in a principal-components analysis[15] of an ensemble of images of natural scenes. In addition, similar values of orientation-tuning half-widths have been obtained in visual search measurements[16] and in single-cell recordings from primate cortex.[17–19]

Although yielding plausible orientation-tuning curves, the analysis[7] was based on some simplifying assumptions, including the rapid convergence of a series expansion of the hypothesized orientation-tuning function and the small size of the increments $\Delta\theta$. More generally, it was not feasible to apply analytic methods to the problem of spatially characterizing the underlying visual filters and their responses to gray-level images. A more direct approach with numerical methods was required.

3. MODEL DESIGN

Models of visual texture segmentation have been used to predict target detection. The tasks are indeed related,[6] but texture-segmentation models have generally been complex, with one,[20] for example, being based on convolution of the image with spatial-frequency- and orientation-selective (Gabor) filters, smoothing by Gaussian filters, thresholding, and final edge detection, and another[21] being based on convolution of the image with even-symmetric linear filters, half-wave rectification, a second nonlinear transformation, and texture-boundary detection. The present model was intended to be simple and to apply only to line-element detection; its division into two image-processing stages was based on a standard approach[22] and, within each stage, the number of free parameters was minimized subject to the requirement that the model accounted quantitatively for the psychophysical data.

4. GENERATION OF TEST IMAGES

Each line element was generated as a gray-level image (0–255 levels) within a region of size 64 × 64 pixels. The line element, which had variable orientation, was positioned at the center of the region, and its dimensions were fixed at 32 × 4 pixels, with the length/width ratio being chosen so as to be similar to that of the line elements used psychophysically, despite its having been generated by a calligraphic rather than a raster-graphic system.[7]

Two methods of producing oriented line elements were considered. The first method used antialiasing to eliminate staircasing that would otherwise be present for all the line elements at orientations other than the vertical and the horizontal; it allowed the full gray-level range to be exploited, and it produced smooth edges. In practice this method proved unsuitable because the antialiasing introduced variability in the energy statistics of the line elements according to their orientation. The second method simply generated bit images in which pixels took on one of only two gray levels, 0 and 255. Although this method suffered from staircasing, line elements had identical energy statistics, and the latter was considered to be the more important factor.

Oriented line elements were thus produced at 5°-orientation intervals. Test images were then composed as a collage of these 64 × 64 blocks (compare Fig. 1).

5. STAGE 1. LINEAR ORIENTED FILTERS

In the first stage of the model, differences in line-element orientation were converted into differences in local response amplitude. The linear neighborhood filter functions f were given the form of a Gaussian function in one spatial direction, the preferred direction, multiplied by a difference-of-Gaussians function in the orthogonal direction. Thus, in the frontoparallel plane with the usual xy-coordinate system, if the preferred orientation were along the x axis, then

$$f(x, y) = \exp\left(\frac{-x^2}{2\sigma_1^2}\right)\left[\exp\left(\frac{-y^2}{2\sigma_2^2}\right) - \frac{\sigma_2}{\sigma_3}\exp\left(\frac{-y^2}{2\sigma_3^2}\right)\right],$$

$$(1)$$

where σ_i, $i = 1, 2, 3$, are the standard deviations (space constants) of the Gaussian functions. Figure 3 shows the general form of the filter. The difference-of-Gaussians function provided balanced excitatory and inhibitory activity in the direction orthogonal to the preferred direction and therefore favored energy distributions oriented along the preferred direction.

Let f_ϕ denote the filter f with a preferred orientation ϕ. Two classes of filters f_ϕ were produced according to Eq. (1), with preferred orientations $\phi = \phi_v$, ϕ_h near the vertical and the horizontal, respectively. In principle, other anisotropic filters could have been used, for example, Gabor functions and difference-of-offset-Gaussians functions, depending on the level of biological plausibility desired and the type of test image used.[21,23] In exploratory simulations poorer predicted responses to the test images were obtained with difference-of-offset-Gaussians functions than with difference-of-Gaussians functions.

The action of a spatially continuous two-dimensional array of these filters on an image was equivalent to a two-dimensional convolution $*$; that is, if the test image had intensity $i(x, y)$ at (x, y), then the response $r_\phi(x, y)$ of each class of filter f_ϕ, $\phi = \phi_v$, ϕ_h, at (x, y) was given by

$$r_\phi(x, y) = (f_\phi * i)(x, y). \tag{2}$$

These responses were processed separately in vertical and horizontal channels defined by the values of ϕ_v, ϕ_h. Independent of the value of ϕ, the response r_ϕ to a uniform field was zero.

6. STAGE 2. NONLINEAR POINT TRANSFORMATION

In the second stage of the model, differences in local response amplitude were converted into differences in local response energy. The responses r_ϕ defined by Eq. (2) were subjected to a power transformation with exponent $\gamma > 0$ and were then given a variable offset k to control the resting level of activity. Because the qualitative behavior of the model was found to be moderately insensitive to the value of γ, its value was fixed at 2, thereby representing a proper conversion of amplitude to energy. The transformed response $r'_\phi(x, y)$ in each channel at each point (x, y) was therefore given by

$$r'_\phi(x, y) = |r(x, y)|^2 + k. \tag{3}$$

7. STAGE 3. SIGNAL-TO-NOISE CALCULATION

Some method of combining filter responses was necessary. The region over which integration takes place need not, however, be the whole image field (for example, within the one image it should be possible for a region containing a target to be compared with another region not containing a target); nor need the integration increase linearly with the number of elements in the region. The target and the nontarget images considered here were presented separately, and responses were therefore integrated over the whole field. A control simulation showed that reducing the number of nontarget line elements did not have a strong effect on the pattern of performance.

The measure of detectability of the line-element target was based on a signal-to-noise calculation for each of the vertical and the horizontal channels: the signal was the total activity over x, y defined by the responses $r'_\phi(x, y)$ to a target image containing $n - 1$ line elements of orientation θ and one line element of orientation $\theta + \Delta\theta$; the noise was the total activity over x, y defined by the responses $r'_\phi(x, y)$ to a nontarget image containing n line elements of orientation θ. The largest signal-to-noise ratio over channels $\phi = \phi_v$, ϕ_h was taken to define a machine discrimination index d'_m. (Strictly speaking, this index may be only monotonically related to a true d': the exponent γ of the nonlinear point transformation is not unique, and, in the absence of a target when performance should be at a chance level, the index is unity, not zero. For ease of interpretation the value of d'_m was floored by subtraction of 1.0, but the adjustment had little effect because the scale and the zero offset of d'_m were treated as free variables in the optimization.) Thus

$$d'_m = \max_{\phi = \phi_v, \phi_h} \left\{ \frac{(n - 1)R(\theta, \phi) + R(\theta + \Delta\theta, \phi)}{nR(\theta, \phi)} \right\} - 1.0. \tag{4}$$

The two classes of filters associated with the two channels had identical spatial-tuning properties and therefore the same gains.

The form of Eq. (4) implies that discrimination is invariant to luminance contrast in the image: the response of the filters to a uniform field is zero (Section 5), and, because the offset k in Eq. (3) is a free variable, any scaling of the intensity $i(x, y)$ at each (x, y) by a constant factor disappears in forming the quotient of the integrated responses.

Notice that discrimination performance is determined by the relative responses of filters to target and nontarget line elements rather than by their absolute responses. Thus detection performance for a vertical target among tilted line elements would be determined by the horizontal channel, even though the response of this channel to the target is weaker than the response of the vertical channel. (It is this property that leads to the asymmetry in the predicted dependence of $\Delta\theta$ on θ near the vertical and the
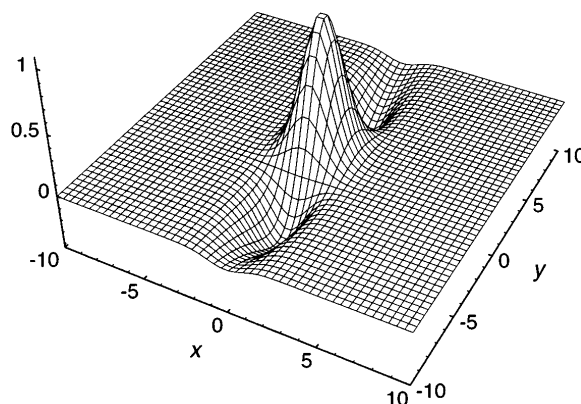
Fig. 3. General form of the anisotropic linear filters.

horizontal.) The dominance of relative responses over absolute responses has been suggested as an explanation for the fineness of orientation discrimination in some acuity tasks.[24,25]

8. FILTER OPTIMIZATION

For a given set of parameter values the model straightforwardly yielded an increment-threshold function for orientation, as follows. For each background orientation θ, values of the machine discrimination index d'_m [Eq. (4)] were computed for progressively increasing values of the relative orientation $\Delta\theta$ of the target element. From the resulting machine psychometric function a threshold value of $\Delta\theta$ was calculated at that θ value for some criterion value of d'_m. The converse problem—to determine the parameter values yielding the given increment-threshold function—was more difficult. The method of simulated annealing[26] used to solve the problem is an optimization procedure that typically has been applied to problems of combinatorial minimization, where there is an objective function to be minimized and where the space over which that function is defined is a discrete, but very large, N-dimensional configuration space. Normally the number of elements in the configuration space is factoriallly large and cannot be explored exhaustively. Although the configuration space was continuous here, simulated annealing could still be applied.[27]

Simulated annealing is, formally, a Monte Carlo iterative improvement method. It is analogous to the way in which a metal is slowly cooled (annealed) so that, given sufficient time, its atoms settle into the lowest possible energy state. Some other optimization algorithms, such as gradient descent, correspond to rapid cooling of the metal. Exploratory computations undertaken with gradient descent showed, not unexpectedly, that parameter values were likely to be trapped in regions corresponding to non-global minima. This problem could have been solved by multiple applications of the gradient-descent operation, each time from a different starting point; such an approach would have eventually yielded a global minimum but also would have been computationally expensive. Simulated annealing offered a more efficient method for covering the parameter space.

As with any optimization procedure, a function quantifying the performance of the model had to be defined first; here it was simply the difference between the increment-threshold function provided by the model and that obtained psychophysically. The calculation was as follows. At each background orientation $\theta = 0°, 22.5°, \ldots, 157.5°$ (corresponding to the eight intervals used to plot the psychophysical data), the machine discrimination index d'_m [Eq. (4)] was evaluated for relative target orientations $\Delta\theta = 0°, 10°, \ldots, 50°$. A criterion value of d'_m was selected (at random because the response of the model could be arbitrarily scaled), and a threshold value of $\Delta\theta$ was calculated. Let $\tau_m(j)$, $\tau(j)$, $j = 1, 2, \ldots, 8$, be the machine and the psychometric increment-threshold values, respectively, at the background values $\theta = 0°, 22.5°, \ldots, 157.5°$. An error function E was computed as the sum of the absolute values of the differences between machine and psychometric increment-threshold values (the latter at $d' = 0.2$). Thus,

$$E = \sum_{j=1}^{8} |\tau(j) - \tau_m(j)|. \tag{5}$$

In all, there were seven parameter values to be varied: the preferred orientations ϕ_v, ϕ_h of the two classes of filters; the space constants σ_i, $i = 1, 2, 3$, of the filters [see Eq. (1)]; the response offset k [Eq. (3)]; and the criterion value of the machine discrimination index d'_m. All the parameter values other than the preferred orientations were common to both classes of filters. The degrees of freedom for the model were therefore one less than for the psychophysical data (the relative gains of the two filter classes also could have been varied,[7] with the consequent loss of the last degree of freedom). This estimate of the degrees of freedom may have been unnecessarily pessimistic because not all the parameters of the model were fully independent of one another.

At the beginning of the optimization procedure the preferred orientation of one filter was set to 0° and the other to 90°; the other parameters were assigned arbitrary values within certain broad ranges. To improve the efficiency of the annealing, a preliminary sequence of iterations was performed in which parameter ranges were adjusted to ensure approximately equal sensitivities of the model to variations in each of the parameter values[28]; the ranges of preferred orientations were limited to $[-5°, 5°]$ and $[85°, 95°]$. Each full annealing sequence comprised 1000 iterations. As a control on its stability the annealing was repeated 50 times, with different seeds for the random-number generator on each occasion.

9. MODEL PERFORMANCE

The optimized model produced an orientation increment-threshold function that closely matched the psychophysical data. The result of the annealing that gave the best fit is shown by the solid curve in Fig. 2; there was no significant difference between the psychophysical and the model data: a test with adequate power based on the chi-squared distribution yielded $\chi^2(1) = 1.0$, $P > 0.3$. Table 1 gives summary statistics for the 50 repetitions. The mean and the standard deviations of the parameters are shown, but only for the filter characteristics, because the value of the offset k and the criterion value of d'_m were only of technical interest. More detailed statistical analysis of each of the sets of optimized parameter values for the preferred orientations ϕ_v, ϕ_h revealed closely spaced bimodal rather than unimodal distributions, again indicating the complex nature of the multidimensional error space.

Table 1. Summary Statistics for a Sample of 50 Repetitions of the Optimization Procedure

Parameter	Mean[a]	SD[b]
σ_1	15.3	3.7
σ_2	1.76	0.42
σ_3	22.3	7.67
ϕ_v	3.83	1.87
ϕ_h	90.6	3.42

[a]Weighted by the error function E [Eq. (5)]; weighted and unweighted means were similar.
[b]Sample standard deviation; also weighted by E.

10. COMMENT

The results of the optimization showed that a simple three-stage model with two classes of orientation-sensitive filters could account quantitatively for the anisotropies in the psychophysical line-element-detection data.

The preferred orientations ϕ_v and ϕ_h of the optimized filters were, when averaged over the 50 annealings, $3.8 \pm 1.9°$ and $90.6 \pm 3.4°$, respectively (Table 1). The averaged space constants of the optimized filters were such that responses to image energy distributions oriented orthogonally to the preferred orientations of the filters were inhibited over extended areas: the ratio of inhibitory to excitatory space constants was approximately 13:1 (Table 1). One interpretation of the high ratio of inhibitory to excitatory space constants is that the model produces the performance shown in Fig. 2 by discriminating against nonpreferred line elements rather than by favoring preferred line elements. (Previously published estimates[13] of the parameter values were derived for a model with no response offset k and were based on a single annealing. Although these earlier estimates have some properties in common with the mean values reported here and produce a good fit to the psychophysical data, they are significantly different numerically.)

The quantitative performance of the model lends support to the assumption[7,29] that the early visual processes determining line-element-detection performance are dominated by two classes of orientation-sensitive filters with axes near the vertical and the horizontal. Even so, it is not suggested here or elsewhere[7] that these are the only filter mechanisms involved in preattentive line-element detection. Results from a study of categorization in the visual search for oriented line elements[30] have suggested processes sensitive to at least two orientations (steep, shallow) and two directions (tilted left, tilted right). Furthermore, a statistical analysis[31] of line-element-detection performance with orientation intervals of 5° rather than 22.5°, as considered here, has revealed fine variations in orientation thresholds superimposed on the broad bimodal performance shown in Fig. 2. This fine orientation structure is likely to be due to the action of additional classes of orientation-sensitive filters, which appear on the basis of psychophysical measurements to make an increasing contribution to detection as stimulus duration increases.[31] In principle, the model considered here could be extended to include these additional filter classes.

ACKNOWLEDGMENTS

REFERENCES

1. J. Beck and B. Ambler, "The effects of concentrated and distributed attention on peripheral acuity," Percept. Psychophys. **14,** 225–230 (1973).
2. J. R. Bergen and B. Julesz, "Parallel versus serial processing in rapid pattern discrimination," Nature (London) **303,** 696–698 (1983).
3. A. Treisman, "Preattentive processing in vision," Comput. Vision Graphics Image Process. **31,** 156–177 (1985).
4. D. Sagi and B. Julesz, "'Where' and 'what' in vision," Science **228,** 1217–1219 (1985).
5. A. Javadnia and K. H. Ruddock, "The limits of parallel processing in the visual discrimination of orientation and magnification," Spatial Vision **3,** 97–114 (1988).
6. H. C. Nothdurft, "Texture segmentation and pop-out from orientation contrast," Vision Res. **31,** 1073–1078 (1991).
7. D. H. Foster and P. A. Ward, "Asymmetries in oriented-line detection indicate two orthogonal filters in early vision," Proc. R. Soc. London Ser. B **243,** 75–81 (1991).
8. J. M. Wolfe and S. R. Friedman-Hill, "Visual search for oriented lines: the role of angular relations between targets and distractors," Spatial Vision **6,** 199–207 (1992).
9. T. Meigen, W.-D. Lagrèze, and M. Bach, "Asymmetries in preattentive line detection," Vision Res. **34,** 3103–3109 (1994).
10. C. Marendaz, P. Stivalet, L. Barraclough, and P. Walkowiac, "Effect of gravitational cues on visual search for orientation," J. Exp. Psychol. Hum. Percept. Perform. **19,** 1266–1277 (1993).
11. J. Beck and B. Ambler, "Discriminability of differences in line slope and in line arrangement as a function of mask delay," Percept. Psychophys. **12,** 33–38 (1972).
12. H.-C. Nothdurft, "Feature analysis and the role of similarity in preattentive vision," Percept. Psychophys. **52,** 355–375 (1992).
13. S. Westland and D. H. Foster, "A line-target detection model using horizontal-vertical filters," in *Visual Search III,* A. Gale, K. Carr, and I. R. Moorhead, eds. (Taylor & Francis, London, to be published).
14. D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).
15. R. J. Baddeley and P. J. B. Hancock, "A statistical analysis of natural images matches psychophysically derived orientation tuning curves," Proc. R. Soc. London Ser. B **246,** 219–223 (1991).
16. W. F. Alkhateeb, R. J. Morris, and K. H. Ruddock, "Effects of stimulus complexity on simple spatial discriminations," Spatial Vision **5,** 129–141 (1990).
17. P. H. Schiller, B. L. Finlay, and S. F. Volman, "Quantitative studies of single-cell properties in monkey striate cortex. II. Orientation specificity and ocular dominance," J. Neurophysiol. **39,** 1320–1333 (1976).
18. R. L. De Valois, D. G. Albrecht, and L. G. Thorell, "Spatial frequency selectivity of cells in macaque visual cortex," Vision Res. **22,** 545–559 (1982).
19. R. L. De Valois, E. W. Yund, and N. Hepler, "The orientation and direction selectivity of cells in macaque visual cortex," Vision Res. **22,** 531–544 (1982).
20. B. S. Rubenstein and D. Sagi, "Spatial variability as a limiting factor in texture-discrimination tasks: implications for performance asymmetries," J. Opt. Soc. Am. A **7,** 1632–1643 (1990).
21. J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms," J. Opt. Soc. Am. A **7,** 923–932 (1990).
22. M. Unser and M. Eden, "Nonlinear operators for improving texture segmentation based on features extracted by spatial filtering," IEEE Trans. Syst. Man Cybern. **20,** 804–815 (1990).
23. R. A. Young, "The Gaussian derivative model for spatial vision. I. Retinal mechanisms," Spatial Vision **2,** 273–293 (1987).
24. G. Westheimer, K. Shimamura, and S. P. McKee, "Interference with line-orientation sensitivity," J. Opt. Soc. Am. **66,** 332–338 (1976).
25. D. Regan and K. I. Beverley, "Postadaptation orientation discrimination," J. Opt. Soc. Am. A **2,** 147–155 (1985).
26. J. Pannetier, "Simulated annealing: an introductory review," in *Neutron Scattering Data Analysis,* M. Johnson, ed. (Institute of Physics, Bristol, UK, 1990), pp. 23–44.
27. W. H. Press and S. A. Teukolsky, "Simulated annealing optimization over continuous spaces," Comput. Phys. July/Aug., 426–429 (1991).
28. L. Ingber and B. Rosen, "Genetic algorithms and very fast

simulated reannealing: a comparison," Math. Comput. Model. **16,** 87–100 (1992).

29. D. H. Foster and P. A. Ward, "Horizontal–vertical filters in early vision predict anomalous line-orientation identification frequencies," Proc. R. Soc. London Ser. B **243,** 83–86 (1991).

30. J. M. Wolfe, S. R. Friedman-Hill, M. I. Stewart, and K. M. O'Connell, "The role of categorization in visual search for orientation," J. Exp. Psychol. Hum. Percept. Perform. **18,** 34–49 (1992).

31. D. H. Foster and S. Westland, "Orientation fine-structure in line-target detection," in *Visual Search III,* A. Gale, K. Carr, and I. R. Moorhead, eds. (Taylor & Francis, London, to be published).