



# Edge co-occurrence in natural images predicts contour grouping performance

W.S. Geisler <sup>a,\*</sup>, J.S. Perry <sup>a</sup>, B.J. Super <sup>b</sup>, D.P. Gallogly <sup>a</sup>

<sup>a</sup> Department of Psychology, University of Texas at Austin, Austin, TX 78712, USA

<sup>b</sup> Department of Electrical Engineering and Computer Science, University of Illinois at Chicago, Chicago, IL 60607, USA

Received 29 March 2000; received in revised form 13 October 2000

## Abstract

The human brain manages to correctly interpret almost every visual image it receives from the environment. Underlying this ability are contour grouping mechanisms that appropriately link local edge elements into global contours. Although a general view of how the brain achieves effective contour grouping has emerged, there have been a number of different specific proposals and few successes at quantitatively predicting performance. These previous proposals have been developed largely by intuition and computational trial and error. A more principled approach is to begin with an examination of the statistical properties of contours that exist in natural images, because it is these statistics that drove the evolution of the grouping mechanisms. Here we report measurements of both absolute and Bayesian edge co-occurrence statistics in natural images, as well as human performance for detecting natural-shaped contours in complex backgrounds. We find that contour detection performance is quantitatively predicted by a local grouping rule derived directly from the co-occurrence statistics, in combination with a very simple integration rule (a transitivity rule) that links the locally grouped contour elements into longer contours. © 2001 Elsevier Science Ltd. All rights reserved.

*Keywords:* Contour perception; Form perception; Grouping; Natural images

## 1. Introduction

During the past decade, a number of psychophysical (Kellman & Shipley, 1991; Field, Hayes, & Hess, 1993; McIlhagga & Mullen, 1996; Dakin & Hess, 1998), neurophysiological (Bosking, Zhang, Schofield, & Fitzpatrick, 1997; Kapadia, Westheimer, & Gilbert, 1999), and computational (Sha'ashua & Ullman, 1988; Parent & Zucker, 1989; Yen & Finkel, 1998) studies have helped develop the 75-year-old Gestalt notion of 'good continuation' (Wertheimer, 1958) into rigorous accounts of contour grouping. The general view that has emerged can be summarized using three processing stages. In the first stage, local contour elements are extracted using spatial filtering like that observed in primary visual cortex. In the second stage, 'binding strengths' are formed between local contour elements. These binding strengths are often described by a *local*

*grouping function* (an 'association field' according to Field et al., 1993), which specifies binding strength as a function of differences in the position, orientation, contrast, and so on, of the contour elements. In the third stage, an integration process uses the local binding strengths to group the local elements into global contours.

If the grouping mechanisms evolved to optimize contour perception in the natural environment then the shape of the local grouping function should be closely related to the statistical co-occurrence of edge elements in natural images, and explanations of contour grouping that are based upon edge co-occurrence should better account for human ability to detect natural contours in complex backgrounds. To test these predictions we measured the probabilities of all possible geometrical relationships between pairs of edge elements extracted from natural images. We measured both the *absolute co-occurrence statistics*, which in principle could be 'learned' directly from the images without feedback, and the *Bayesian co-occurrence statistics*,

\* Corresponding author. Fax: +1-512-4717356.

E-mail address: geisler@psy.utexas.edu (W.S. Geisler).

which would require feedback via interaction with the environment.<sup>1</sup> Then, we derived a local grouping function directly from the edge co-occurrence probabilities, combined this local grouping function with a very simple integration rule, and compared the predictions to parametric detection data for naturalistic contours.

## 2. Absolute edge co-occurrence statistics

### 2.1. Methods

#### 2.1.1. Edge extraction

The first step in measuring the edge co-occurrence statistics is to extract edge elements from natural images. Edge elements were extracted from 20 representative natural images by measuring local contrast energy across orientation at all potential edge locations. Twenty images were selected to represent a wide range



Fig. 1. Representative natural images. Edge elements were extracted from these 20 images in order to estimate the edge co-occurrence statistics for natural images. The images were obtained from the Kodak web site: <http://www.kodak.com/digitalImaging/samples/imageIntro.shtml>.

<sup>1</sup> The term 'absolute' is used to distinguish simple unconditional probability from conditional probability (Feller, 1968); the absolute co-occurrence statistics are based upon absolute probabilities and the Bayesian co-occurrence statistics are based upon conditional probabilities.

of natural terrain (see Fig. 1). Each image was converted to 8-bit gray scale using Adobe PhotoShop, and then windowed with a circular aperture to a diameter of 480 pixels.

Significant edge elements were then extracted from each image using a method that mimics the response characteristics of neurons in primary visual cortex. First, potential edge locations were identified by filtering the image with a non-oriented log Gabor function having a spatial-frequency bandwidth of 1.5 octaves, and a peak spatial frequency of 0.1 cycles/pixel.<sup>2</sup> Each zero-crossing pixel in the filtered image (within a radius of 216 pixels of the image center) was regarded as a potential edge element.

Second, the original image was filtered with oriented log Gabor functions having a spatial frequency bandwidth of 1.5 octaves and an orientation bandwidth of 40°, which are the average values reported for primate visual cortex (Geisler & Albrecht, 1997). The peak spatial frequency of the log Gabor filters was set to 0.1 cycles/pixel. This value was picked so that the spatial scale of the filter kernels ('receptive fields') were matched to the size of the contour elements used in the psychophysical experiments. Filtered images were obtained for log Gabor functions in both sine and cosine phase, at every 10° of orientation. The sine and cosine filtered images were squared and summed to obtain the contrast energy at each orientation for each zero-crossing pixel. The normalized contrast energy at each orientation was then obtained by dividing the contrast energy at that orientation by the sum of the contrast energies across all orientations, plus a constant. (The constant was set so that the half-saturation contrast of the response to an optimal sine wave grating corresponds to the average in monkey primary visual cortex).

Third, a zero-crossing pixel was regarded as a significant edge element if the peak of the normalized contrast energy across orientation exceeded 10% of the maximum possible response; however, the specific value of this threshold had a relatively minor effect on the results. The orientation of this significant edge element was then obtained by finding the centroid of the contrast energy distribution across orientation. The mathematical details of the edge extraction are given in the Section A.1.

Fig. 3A shows the edge elements extracted from one of the 20 images. Each red pixel shows the location of the center of a significant edge element. The orientation of the edge element is not shown, but we have found that the orientations generally correspond closely to subjective judgments of contour orientation.

<sup>2</sup> In analyzing the natural images we assumed a viewing distance such that the pixel size was the same as in our psychophysical experiments (0.03° on each side). Thus, 0.1 cycles/pixel corresponded to 3.3 cycles deg<sup>-1</sup>.

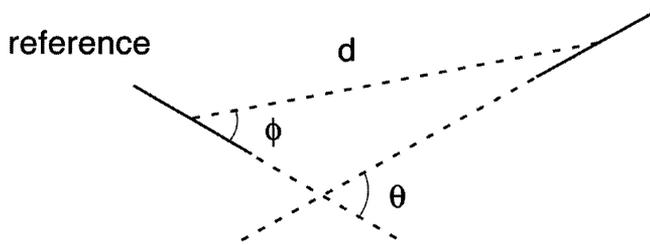


Fig. 2. The geometrical relationship between edge elements. There are three parameters: the distance between the centers of the elements,  $d$ , the orientation difference between the elements,  $\theta$ , and the direction of the second element with respect to the orientation of the reference element,  $\phi$ .

### 2.1.2. Edge co-occurrence probability

The geometrical relationship between any given pair of edge elements can be described with three parameters: the distance between the element centers,  $d$ , the orientation difference between the elements,  $\theta$ , and the direction of the second element relative to the orientation of first (reference) element,  $\phi$  (see Fig. 2). For each image, we compared every edge element with every other edge element, giving us a very large set of triplets ( $d, \theta, \phi$ ), which were then binned into a three-dimensional histogram (bin widths:  $w_d = 8$  pixels,  $w_\theta = 10^\circ$ , and  $w_\phi = 10^\circ$ ).<sup>3</sup> We combined the histograms for all 20 images, and then corrected for the variation in bin size and the effects of the finite circular image aperture (see Section A.2). The corrected histogram gives the absolute probability,  $p(d, \theta, \phi)$ , of observing an edge element at every possible distance, orientation difference, and direction from a given (reference) edge element.<sup>4</sup>

## 2.2. Results

The result of the above statistical analysis is a three-dimensional probability density function,  $p(d, \theta, \phi)$ , derived from the 20 images in Fig. 1. Examination of this three-dimensional function reveals two important statistical properties of natural images. These properties are shown in Fig. 3B and Fig. 3C, which summarize the full probability function. The layout of these plots corresponds directly to the geometry in Fig. 2, with the center horizontal line segment representing the reference element. The radial distance of an element from the reference represents the distance  $d$ , the angular location around the circle (the azimuth) represents the direction  $\phi$ , and the tilt of an element represents the orientation difference  $\theta$ .

<sup>3</sup> Each element in a given pair served as the reference element thus, each pair of elements entered into the histogram twice.

<sup>4</sup> For notational simplicity we let  $(d, \theta, \phi)$  represent both a real valued vector and a discrete valued bin. It should be clear from the context which meaning is intended.

The line segments in Fig. 3B show, at each distance, the most probable orientation difference for each possible direction. In other words, for each of the 6 distances and 36 directions, we located the orientation difference with the highest probability and plotted, at the given direction, a line segment with this orientation difference. The color of the line segment indicates the relative probability.<sup>5</sup> As can be seen, for all distances and all directions, the most likely edge element is approximately parallel to the reference element, with greater probabilities for elements that are nearby and co-linear with the reference element. This result shows that there is a great deal of parallel structure in natural images, presumably due to the effects of growth and erosion (e.g. the parallel sides of a branch, parallel geological strata, etc.), perspective projection (e.g. the elongation of surface markings due to slant), shading (e.g. the illumination of a branch often produces a shading contour parallel to the side of the branch), and so on.

The line segments in Fig. 3C show, at each distance, the most probable direction for each possible orientation difference. In other words, for each of the 6 distances and 36 orientation differences, we located the direction with the highest probability and plotted a line segment with the given orientation difference (tilt) at the bin location of this most probable direction. As can be seen, the two horizontal wedge-shaped regions represent the most likely co-occurring edge elements. Specifically, for all distances and all moderate differences in orientation, the most likely edge element is approximately co-circular with the reference element (i.e., the two elements are tangent to the same circle), with greater probabilities for elements that are nearby and co-linear with the reference element. This result presumably reflects the relatively smooth shapes of natural contours, and it provides direct evidence that the Gestalt principle of good continuation has a general physical basis in the statistics of the natural world.

The results in Fig. 3B and Fig. 3C are quite robust. All of the individual images in our sample show the same basic pattern, and the results for a random selection of 10 images are very similar to those for the remaining 10 images. As a control, histograms were also computed for white noise images; as expected, the resulting plots were random except for a small effect near the reference line due to the effect of the band-pass filtering. Also, a preliminary analysis at a smaller scale (2 octaves higher) suggests that the edge co-occurrence statistics are very similar across spatial scales.

<sup>5</sup> To simplify the probability scale, the three-dimensional probability density function was normalized so that the peak of the function is 1.0.

### 3. Bayesian edge co-occurrence statistics

#### 3.1. Methods

Both evolution and developmental mechanisms could exploit the statistical property in Fig. 3C to form a local grouping function that is qualitatively appropriate for contour grouping (see later). However, for this local grouping function to be computationally appropriate, it would need to be the case that the greater the absolute co-occurrence probability, the more likely the edge elements belong to the same physical contour. Although this is plausible, a local grouping function could be derived from natural image statistics in a more rational fashion using Bayesian statistical decision theory. In the (optimal) Bayesian approach, two edge elements should be bound together if it is more likely that they arose from the same physical contour,  $C$ , than from different physical contours,  $\sim C$ . That is, two elements should be bound together if the likelihood ratio,

$$L(d, \theta, \phi) = \frac{p(d, \theta, \phi | C)}{p(d, \theta, \phi | \sim C)}, \quad (1)$$

is greater than a criterion  $\beta$ , where  $\beta$  is the ratio of the prior probabilities,

$$\beta = \frac{p(\sim C)}{p(C)} = \frac{1 - p(C)}{p(C)}. \quad (2)$$

These Bayesian statistics (the likelihood ratio function and ratio of the prior probabilities) were measured for each image, from the same set of edge elements that were used to measure the absolute co-occurrence statistics. To do this we used an image tracing method (see also Brunswick & Kamiya, 1953; Elder & Goldberg, 1998). Recall that each edge element was centered on a zero-crossing pixel. All of these significant zero-crossing pixels were highlighted in a bright color in the original image (e.g. Fig. 3A). Using a mouse, observers selected zero-crossing pixels that belonged to the same physical contour. Once the observer had selected all the pixels belonging to the same physical contour, those pixels were restored to their values in the original image, and the observer started on a new contour. This procedure continued until all the edge elements were assigned to a unique contour. We found that the observers could process the images with the fewest numbers of edge elements in 2–3 h, whereas the images with the greatest numbers of edge elements might take 6–8 h. The resulting contour assignment information allowed us to compute two conditional co-occurrence probability functions, one for when the elements belonged to the same contour,  $p(d, \theta, \phi | C)$ , and one for when the elements belonged to different contours,  $p(d, \theta, \phi | \sim C)$ . We then took the ratio of these probability functions to obtain the likelihood ratio function (see Eq. (1)).

In assigning edge elements to a contour, the observers were instructed that the same physical contour could be a surface boundary, a surface marking boundary, a shadow boundary or a lighting boundary (i.e. these were to be regarded as different physical sources). To deal with junctions, the observers were instructed not to group across junctions making sharp acute angles. These instructions for junctions are somewhat ambiguous, but because relatively few selected contour elements fall across junctions, the specific instructions must have a relatively minor effect on the overall statistics (although this remains to be tested explicitly).

In our software, the observers were given the ability to zoom in and out, toggle the highlighted edge pixels on and off, switch between the grayscale and full color versions of the image, and make corrections. Under these conditions, the two observers (WSG and a naïve observer CAF) generally reported a high degree of confidence in assigning edge pixels to physical contours. Presumably this confidence derives from the fact that humans are very good at correctly interpreting images when the full set of cues (shape, color, lighting, shading, texture) is available, and when the images are natural, so that all of the observer's experience with the natural world can be brought to bear. The agreement between the two observers was quantified by comparing their final likelihood-ratio functions.

The value of the ratio of the priors,  $\beta$ , was also obtained directly from the image tracing data; however, its value depends strongly on the area of the analysis region; specifically, the values of  $\beta$  increases as the area of the analysis region increases. Thus, our approach has been to ignore the value of  $\beta$  in considering models of performance based upon the Bayesian statistics, and instead, leave it as a single free parameter — a global decision criterion — analogous to the decision criterion in signal detection theory (Green & Swets, 1974).

#### 3.2. Results

The central result of the Bayesian analysis is a three-dimensional function that gives the likelihood ratio that any given pair of edge elements belong to the same physical contour versus different physical contours. Fig. 3D plots this full three-dimensional function based upon all 20 images for both observers (note that here, all 36 orientations are plotted at each distance and direction). The individual functions are quite similar. This is shown in Fig. 4, where the log likelihood ratio values for observer CAF are plotted against the log likelihood ratio values for observer WSG. If the likelihood ratio functions are the same then the data should fall on a straight line of slope one through the origin (the solid line). The correlation between the likelihood ratio functions for the two observers is quite high

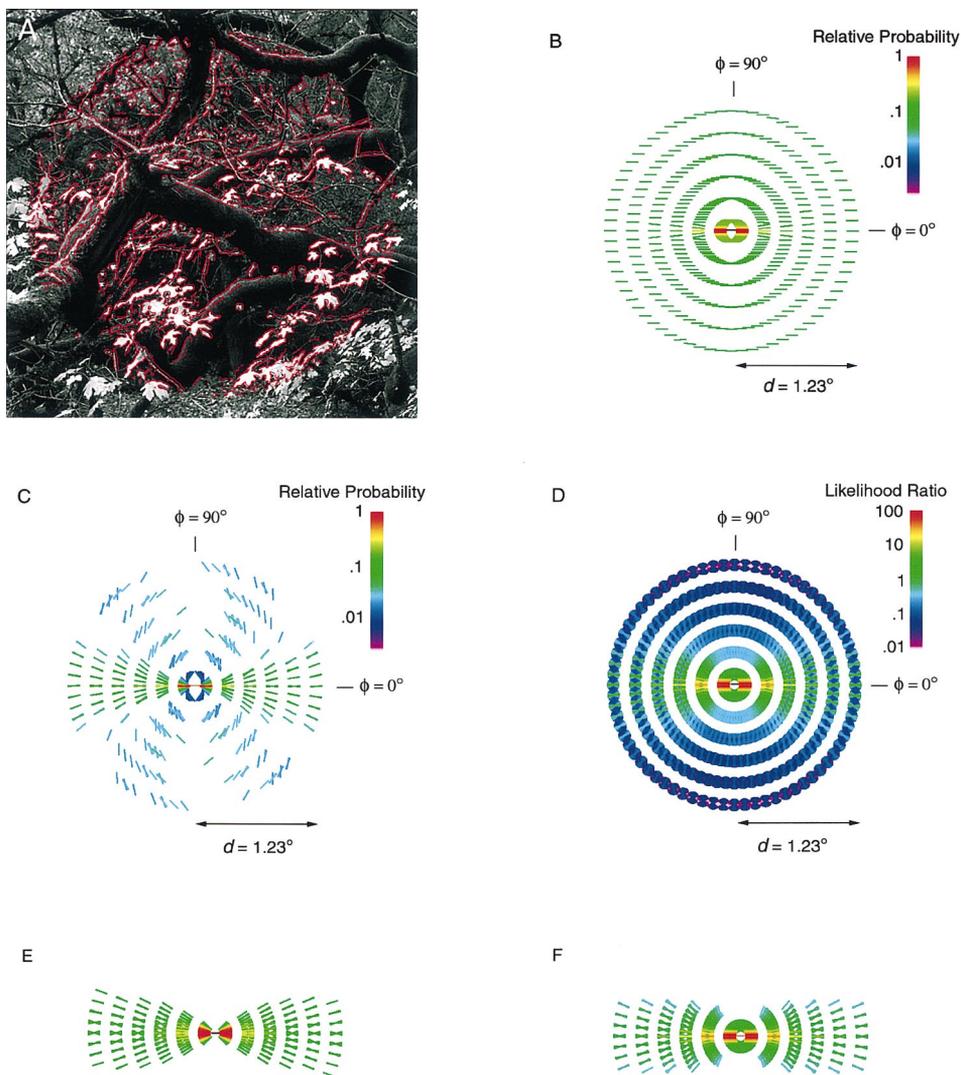


Fig. 3. Statistical analysis of edge co-occurrence in natural images. For each image, edge elements were extracted, and then each edge element was compared with other edge elements. (A) Edge elements. Each red pixel in this image indicates the location of the center of a significant edge element; the orientations of the elements are not shown. (B) First edge co-occurrence property. The line segments show the most frequently occurring orientation difference for each given distance and direction from the central reference element. The color of a line segment indicates the relative probability. (C) Second edge co-occurrence property. The line segments show the most frequently occurring direction for each given distance and orientation difference from the central reference element. (D) Bayesian likelihood-ratio function. Each line segment shows a possible geometrical relationship between an element and the reference element. The color of a line segment indicates the likelihood that the element and reference belong to the same physical contour divided by the likelihood they belong to different physical contours. (E) A thresholded local grouping function derived from the edge co-occurrence plot in C. (F) A thresholded local grouping function derived from the Bayesian likelihood-ratio function in D.

( $r = 0.98$ ). Note that the increased scatter at small log-likelihood values is expected because log probabilities are generally more variable when the estimated probabilities are small.

As can be seen in Fig. 3D, edge elements that are co-circular (i.e. consistent with a smooth continuous contour) are more likely to belong to the same physical contour. These results support our interpretation of the absolute statistics in Fig. 3C, and

provide further evidence that the Gestalt principle of good continuation has a physical basis in the statistics of the natural world. Most importantly, these results allow us to determine a maximum likelihood (optimal) local grouping function for contour grouping in natural scenes. Given the fundamental importance of contour grouping for useful vision, it is possible that the human local grouping function is near this optimum.

## 4. Contour grouping experiment

### 4.1. Methods

In the psychophysical experiments, we measured detection accuracy for line segment contours embedded in backgrounds consisting of randomly oriented line segments (Beck, Rosenfeld, & Ivry, 1989; Field et al., 1993; McIlhagga & Mullen, 1996; Dakin & Hess, 1998).<sup>6</sup> On each trial, the subject viewed one image containing only background elements and another image containing an embedded contour (see Fig. 5A). The two images were presented separately in a random temporal order, and the subject indicated which of the two images contained the contour. After the subject responded, he/she was informed about the correctness of the response and shown the actual locations of the contour elements.

The display had a diameter of  $12^\circ$  (400 pixels at a

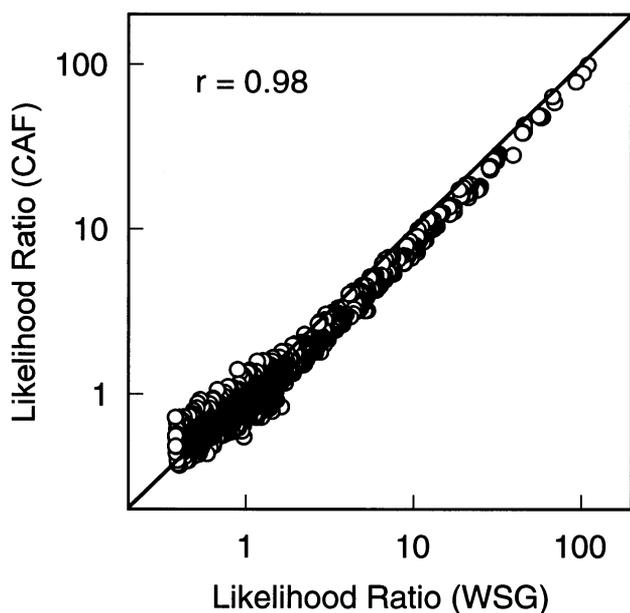


Fig. 4. Comparison of the likelihood ratio functions for observers WSG and CAF. Each point represents a pair of log likelihood-ratio values, one from WSG and one from CAF, for a given bin of the 3D log likelihood-ratio function. The horizontal axis plots the log likelihood-ratio values for WSG and the vertical axis for CAF. If the likelihood ratio functions are the same for both observers then the data points should fall along the diagonal.

<sup>6</sup> The contour elements were line segments rather than the more popular Gabor patch. This choice was dictated by several factors. First, we used line segments in order to maximize the speed of generating the stimuli, so that we could keep the inter-trial interval short in this large (216 condition) survey experiment. Second, most edges in the real world are broadband like line segments, not narrow-band like Gabor patches. Third, the appropriate choice would seem to depend on how the elements are represented at the level where the grouping across the elements takes place. It is quite possible that, at the level where grouping occurs, Gabor patches have the more complex representation (i.e. a complicated 3 bar structure).

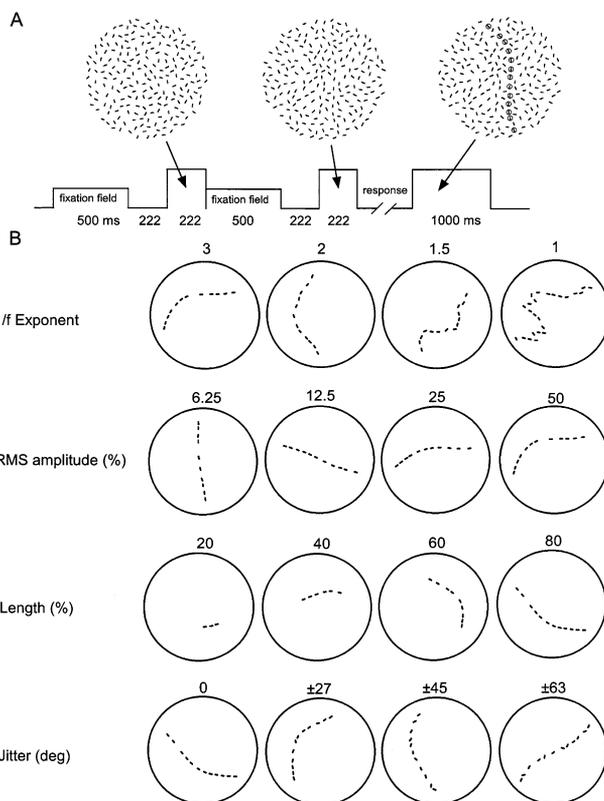


Fig. 5. Stimuli and procedure for the contour grouping experiment. (A) Time line for a single trial of the two-interval, two-alternative forced choice experiment. (B) Examples of target contours for the four stimulus dimensions that were varied factorially in the experiment. These stimulus dimensions were selected to be representative of wide range of contour shapes.

viewing distance of 112 cm), and a uniform green ( $x = 0.29, y = 0.61$ ) luminance of  $64 \text{ cd m}^{-2}$ , except when a fixation, test, or feedback field was present. The fixation field contained a small dark central cross. The test fields contained dark line segments,  $0.3 \times 0.03^\circ$  (note that the elements have been widened in Fig. 5). The target contours were generated by summing sine wave components that modulated about a randomly oriented axis through the center of the display. The spatial frequencies of the components were harmonics of 0.5 cycles per display diameter. A different contour shape was generated on each trial by randomly selecting the amplitudes and phases of the components, and then filtering (i.e. multiplying the amplitudes by a transfer function). Line elements were placed randomly along the contour and then in the background, with the constraints that they could not be closer than half the line element length, and that the density of line elements was the same along the contour and in the background. The contour shapes were made 'naturalistic' by using transfer functions of the form  $1/f^n$  where  $n$  is the '1/f exponent' (Burton & Moorehead, 1987; Field, 1987). Further details of the display generation method are given in Section A.3.

Detection accuracy was measured factorially as a function of four variables: (i) the  $1/f$  exponent (1, 1.5, 2, and 3); (ii) the RMS amplitude of the contour (6.5, 12.5, 25, and 50% of the display diameter); (iii) the contour axis length (20, 40, 60 and 80% of the display radius); and (iv) the range of orientation jitter of the elements ( $0$ ,  $\pm 27$ ,  $\pm 45$ , and  $\pm 63^\circ$ ). Altogether, 216 different classes of random contour shape were tested. Some examples of these contour shapes are shown in Fig. 5B.

The subjects were first practiced in the task until near asymptotic performance was reached. Then, two 30 trial blocks were run for each of the 216 classes of contour shape, for each subject. The order of the conditions was designed to minimize the effects of practice on the overall trends. In addition, the two blocks for each condition were obtained using an ‘ABBA’ procedure; all 216 conditions were run in one order and then run again in the reverse order.

## 4.2. Results

The solid symbols in Fig. 6A–D show the psychophysical results for the two subjects. In agreement with previous studies (Field et al., 1993; McIlhagga & Mullen, 1996; Dakin & Hess, 1998), the data show that humans are quite good at detecting contours embedded in noise even when there is great randomness in the location, orientation, and shape of the target contour. For example, contours like that in the middle picture of Fig. 5A are detected with better than 90% accuracy. As can be seen, performance generally improves with increases in the  $1/f$  exponent and contour length, and generally declines with increases in RMS amplitude and jitter. These experimental results provide a detailed parametric measurement of human ability to detect naturalistic contours in noisy backgrounds.

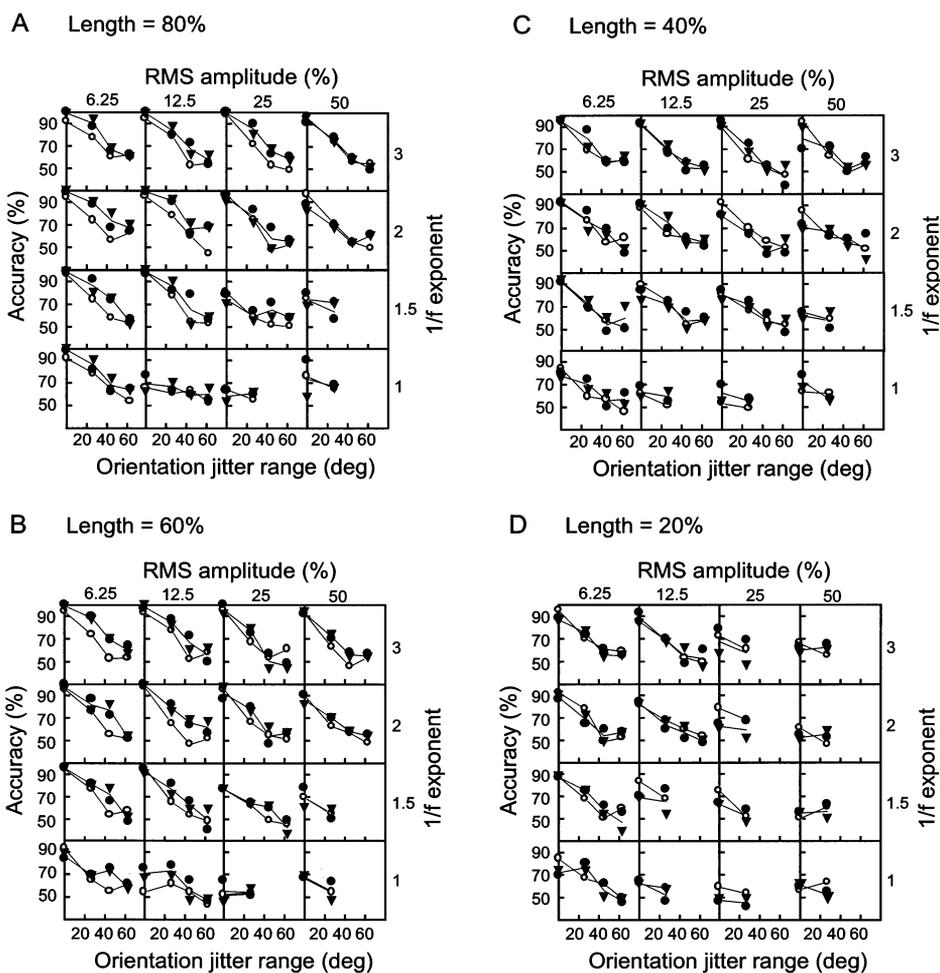


Fig. 6. Results and predictions for the contour grouping experiment. The figure plots detection accuracy in percent correct as a function of the contour shape ( $1/f$  exponent and RMS amplitude), contour length, and the range of orientation jitter of the contour elements. The panels with only two points indicate conditions where the highest levels of jitter were not tested because pilot experiments showed that performance was poor even with 0% jitter. The solid triangles show the data for subject WSG and the solid circles for subject JSP. The open circles show the predictions obtained by applying a local grouping function (based directly upon the edge co-occurrence statistics) followed by a simple transitivity rule: if (edge element)  $a$  binds to  $b$ , and  $b$  binds to  $c$ , then  $a$  binds to  $c$ .

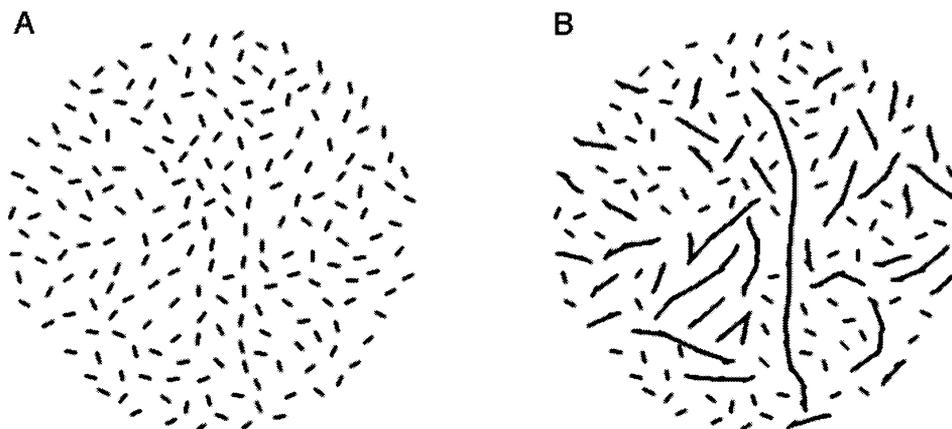


Fig. 7. Predicted contour groups. (A) An example stimulus that contains a target contour. (B) The connected line segments show the groups obtained by applying a local grouping function (based directly upon the edge co-occurrence statistics) followed by a simple transitivity rule: if (edge element)  $a$  binds to  $b$ , and  $b$  binds to  $c$ , then  $a$  becomes bound to  $c$ .

## 5. Predictions

Is the local grouping function that the visual system uses for contour grouping consistent with the co-occurrence statistics of natural images? To address this question, we generated predictions for the psychophysical experiment.

Consider first the absolute co-occurrence statistics. To generate predictions using the edge co-occurrence data in Fig. 3C, it is necessary to introduce two parameters. The first is a ‘tolerance’ parameter,  $\sigma$ , which assigns a relative probability to orientation differences around the maximum probabilities shown in Fig. 3C. This parameter produces a small family of local grouping functions, all of which are consistent with the data in Fig. 3C. The second parameter is a binding criterion,  $\beta$ , which is a threshold on the local grouping function: any pair of edge elements whose co-occurrence probability exceeds this criterion get bound together, otherwise they do not. An example of a thresholded local grouping function is shown in Fig. 3E. The line segments drawn in this diagram show all the specific combinations of distance, orientation difference, and direction that result in grouping to the reference element. Predictions for the contour detection task were obtained by combining a thresholded local grouping function with the simplest possible integration mechanism — a transitivity rule: if (edge element)  $a$  binds to  $b$ , and  $b$  binds to  $c$ , then  $a$  becomes bound to  $c$  (Geisler & Super, 2000).

A predicted response (‘first’ or ‘second’ interval) was computed for each of the specific stimuli presented in each trial of the experiment, for each subject. The predictions for a given trial were generated as follows. First, we retrieved the specific list of line segments presented in the first and second intervals of the trial. Second, each line segment was compared with every

other line segment, and a pair of line segments was bound together if they satisfied one of the geometrical relationships specified in the thresholded local grouping function (i.e. Fig. 3E). Third, the final groups were obtained by applying the transitivity rule over all the two-element groups formed in the second step. Fourth, and finally, we assumed that, in our two-interval forced choice task, observers selected the interval containing the longest group that was consistent with the possible contour locations. The result of this processing is demonstrated in Fig. 7B, where the connected elements show the final groups that are formed for the image in Fig. 7A. As can be seen, the longest group corresponds to the target contour.

The open circles in Fig. 6A–D show the predictions for all the data using the thresholded local grouping function in Fig. 3E, which was obtained from Fig. 3C by setting  $\sigma = 20^\circ$  and  $\beta = 0.05$ . As can be seen, the predictions are remarkably good (given only two free parameters). The value of the Pearson correlation coefficient for the predicted accuracy versus the observed accuracy is 0.87.

To generate predictions using the Bayesian co-occurrence statistics only one free parameter is required — the likelihood-ratio criterion  $\beta$ . Recall that the rational decision rule is to group a given pair of edge elements together if the likelihood ratio exceeds this criterion. Thus, the optimal local grouping function is obtained by thresholding the likelihood ratio histogram shown in Fig. 3D. The resulting function for the best fitting value of  $\beta$  (0.38) is shown in Fig. 3F. The psychophysical predictions are not shown here, but they are slightly better than those obtained using the absolute co-occurrence statistics. The value of the Pearson correlation coefficient for the predicted accuracy versus the observed accuracy is 0.89.

For both the absolute and Bayesian predictions, an inspection of individual displays indicates that there is generally good agreement between the contours one 'sees' and the contours found with the local grouping function and transitivity rule (e.g. see Fig. 7). Further, the parameter values that give the best fit to the human data are also the values that maximize the overall detection accuracy of the model. Finally, we have tested some other forms for the local grouping function and have been unable to find one that produces significantly higher accuracy than the ones based upon the natural-image statistics.

## 6. Discussion

In this study, we measured the absolute and Bayesian (conditional) co-occurrence statistics of edge elements extracted from a representative collection of natural images, and we measured the detectability of naturalistic contours in complex backgrounds for a wide range of contour shapes. The absolute edge co-occurrence statistics of natural images were found to have two fundamental properties, one corresponding to the Gestalt principle of 'good continuation' (Fig. 3C), and one corresponding to the parallel structure of the visual world (Fig. 3B). Further, the Bayesian co-occurrence statistics were found to have a structure similar to the good-continuation property of the absolute statistics (c.f. Fig. 3C and Fig. 3D). Also, we found that the absolute and Bayesian statistics imply similar local grouping functions (c.f. Fig. 3E and Fig. 3F). Finally, and most remarkably, we found that human performance for detecting naturalistic contour shapes in the complex backgrounds is predicted fairly accurately by these local grouping functions, in conjunction with a simple transitivity rule.

### 6.1. Image statistics

The fundamental importance of measuring the statistical properties of images that are relevant to the performance of visual tasks has long been recognized. For example, Brunswick (Brunswick & Kamiya, 1953) and Gibson (1966) make a strong case for conducting ecological analyses of the images arising from the environment. They clearly recognized that it is the lawful structure and statistical correlations of the visual world that allow the visual system to function so amazingly well in interpreting the images formed on the retinas. A related argument has been made for the importance of measuring the information relevant to a visual task using Bayesian ideal-observer theory (e.g. Geisler, 1989; Knill & Richards, 1996), and obviously, it is not possible to develop ideal observers for natural visual tasks without measuring the statistical

properties of the natural images that are relevant for those tasks.

Most recent efforts to measure the statistics of natural images have focused on measuring one-dimensional statistics, such as the spatial frequency spectra of natural images (Burton & Moorehead, 1987; Field, 1987), the orientation spectra of natural images (Switkes, Mayer, & Sloan, 1978; Coppola, Purves, McCoy, & Purves, 1998), the wavelength spectra of natural illuminants (Judd et al., 1964; Dixon, 1978), and the reflectance spectra of natural materials (Buchsbaum & Gottschalk, 1984; Maloney, 1986). From these studies, we now know that natural spatial-frequency spectra fall off as  $1/f^n$ , that the orientation spectra peak in the vertical and horizontal orientations, and that the spectra of natural illuminants and materials can be described with just a few parameters each. These results have proven to be highly useful for understanding the initial encoding of spatial and chromatic information in the visual system, as well as for understanding color constancy.

However, for understanding perceptual grouping, it is necessary to measure multidimensional statistics, such as the absolute and Bayesian co-occurrence statistics reported here. Recently, within the context of image compression, image restoration, and texture synthesis, there have been efforts to examine the absolute co-occurrence statistics of image features (such as the co-occurrence statistics of wavelet coefficients); for a brief summary, see Simoncelli (1997). These statistics have not been measured for a wide range of natural images, or with the aim of measuring the information relevant for contour grouping, and hence are not directly comparable to the current results.

Brunswick and Kamiya (1953) hypothesized that environmental image statistics may underlie the Gestalt principles of perceptual grouping, and indeed they provided some evidence that nearby parallel contours are more likely to arise from the same physical object than are distant parallel contours. They took this result as support for the 'ecological validity' of grouping on the basis of 'proximity'. Without the availability of high speed computers, their analysis was necessarily very limited, but conceptually it is closely related to what we refer to as the measurement of the Bayesian co-occurrence statistics: measurement of the probability distributions needed to determine the likelihood ratios and prior probabilities that particular image features arise from the same physical source. The measurements reported here, and some related measurements made independently by Elder and Goldberg (1998), appear to be the only attempts since Brunswick and Kamiya (1953) to measure the Bayesian (conditional) co-occurrence statistics that are relevant for perceptual grouping.

## 6.2. Absolute versus Bayesian statistics

The present finding, that the local grouping functions derived from the absolute and Bayesian statistics are very similar, may have important implications for understanding the evolution and development of the contour grouping mechanisms, because local grouping functions based upon Bayesian statistics are relatively optimal. While one might expect the slow process of evolution to eventually incorporate the optimal Bayesian co-occurrence statistics into a local grouping function, it would be much more difficult for learning or developmental processes to do so because learning of the Bayesian statistics requires knowing whether or not pairs of edge elements arise from the same physical contour. An advantage of the absolute co-occurrence statistics is that they can be learned directly from the images without reference to the physical world. Simulations have shown that self-organizing networks can respond to, and hence incorporate, co-occurrence statistics (Kohonen, 1989; Miller, Keller, & Stryker, 1989; Sirosh & Miikkulainen, 1994). Thus, the existence of the statistical properties shown in Fig. 3C, and their success in predicting human performance, support the possibility that self-organizing neural mechanisms play an important role in the learning, maintenance and/or shaping of the local grouping functions that underlie contour grouping.

## 6.3. Models of contour grouping

As mentioned earlier, the general framework underlying existing models of contour grouping consists of three processing stages: (1) extraction of local contour elements using spatial filtering like that observed in primary visual cortex; (2) formation of binding strengths between local contour elements according to some form of local grouping function; and (3) integration of local elements into global contours based upon the local binding strengths.

In most models, the extraction of local contour elements is obtained by filtering with spatially oriented 'receptive fields' consisting of the second derivative of a Gaussian (or a difference of Gaussians) in one direction, multiplied by a Gaussian in the orthogonal direction (Sha'ashua & Ullman, 1988; Parent & Zucker, 1989; Gigus & Malik, 1991; Dakin, 1997; Yen & Finkel, 1998; Hess & Dakin, 1999). The specific form of the filtering and orientation estimation is not very important for the qualitative behavior of the models, but can have some quantitative effect. We chose the 'parameter-free' approach of using filter shapes based directly upon the measured average tuning characteristics of neurons in monkey primary visual cortex (e.g. see Geisler & Albrecht, 1997). Another parameter-free approach is to use filter shapes derived from psycho-

physical spatial masking studies (e.g. see Hess & Dakin, 1999). The filter shapes in these two parameter-free approaches are quite similar.

The local grouping functions in most models of contour grouping are based upon three principles: co-circularity, smoothness, and proximity (Sha'ashua & Ullman, 1988; Parent & Zucker, 1989; Gigus & Malik, 1991; Kellman & Shipley, 1991; Dakin, 1997; Pettet, McKee & Grzywacz, 1998; Yen & Finkel, 1998; Hess & Dakin, 1999; Geisler, Thornton Gallogly, & Perry, 2000). The greater is the proximity, co-circularity, and smoothness, the stronger is the local binding strength. (Note that two contour elements are co-circular if they are both tangent to a circle of some radius, and are smooth if the radius of the circle is large.) However, the specific shape of the local grouping function and the way it is implemented differ. Dakin (1997) and Hess and Dakin (1999) do not incorporate the co-circularity constraint. Sha'ashua and Ullman (1988), Parent and Zucker (1989), Yen and Finkel (1998), and Pettet, McKee, and Grzywacz (1998) implement a local grouping function via an iterative feedback process, whereas Gigus and Malik (1991), Dakin (1997), Hess and Dakin (1999) and Geisler et al. (2000) implement a simpler one-step process (although the underlying neural mechanisms may still involve some local feedback). Here, we also use a simple one-step process. However, unlike previous models, the local grouping function proposed here is based directly upon the statistics of natural images, which we find are approximately consistent with the principles of co-circularity, smoothness and proximity.

The final stage, where local elements are integrated into global contours, is only incorporated into some of the models (Sha'ashua & Ullman, 1988; Pettet, McKee & Grzywacz, 1998; Yen & Finkel, 1998; Geisler et al. 2000), and again, the implementations are different. Sha'ashua and Ullman (1988), Yen and Finkel (1998), and Pettet, McKee and Grzywacz (1998) find extended contours by thresholding a 'saliency map,' which is the output from the iterative feed-back second stage; whereas, Geisler et al. (2000), like the present study, apply a parameter-free transitive grouping rule: if (element)  $a$  binds to  $b$  and  $b$  binds to  $c$ , then  $a$  binds to  $c$ .

Geisler and Super (2000) argue that transitive grouping is a general principle that is likely to play a fundamental role in many aspects of perceptual organization, not just contour integration. They note that the effects of perspective projection, natural lighting, and growth and erosion generally result in smooth variations in the local spatial properties of the retinal images of surfaces and surface boundaries, and therefore, transitive grouping mechanisms are essential in order to correctly group image features that belong to the same surface or surface boundary. The specific form of transitive grouping that they propose (the above rule) is very simple,

testable, and has been successful in accounting for grouping data. However, the rule will probably be found to be too simple. Nonetheless, it seems likely that some form of transitive grouping is implemented in the visual system.

A limitation of the present model is that it is deterministic. Contour elements are directly linked only if they satisfy one of the geometric relationships given by the local grouping function (see Fig. 3E and Fig. 3F), and are then indirectly linked by the fixed transitive grouping rule. However, the model can be made probabilistic, with little effect on the fits, by assuming some local or global variance in the value of the binding criterion  $\beta$ .

It is interesting to note that the present model predicts superior performance when the target contours form ‘closed’ as opposed to ‘open’ contours (Kovacs & Julesz, 1993; Pettet, McKee, & Grzywacz, 1998), even though the model does not include an iterative feedback mechanism (reverberating loop). The reason is simply that breaking a closed contour decreases the mean length of the longest group (i.e. long groups that would have crossed the break do not get formed). It remains to be seen whether this predicted effect of closure is large enough to account for the psychophysically measured effect of closure.

#### 6.4. Neural basis of the local grouping function

Presumably, the success of the psychophysical predictions reported here (with almost no free parameters) is not an accident, but reflects the existence of simple neural mechanisms corresponding to the local grouping function and the transitive grouping mechanism. The obvious hypothesis for the local grouping function is a neural population with a receptive field structure matched to the edge co-occurrence statistics. Although the most appropriate experiments have yet to be done, the current evidence suggests that this neural population does not reside in primary visual cortex (for example, see Walker, Ohzawa, & Freeman, 1999), and hence is more likely to be found in later cortical areas. There is even less physiological evidence relevant to the transitive grouping mechanism. However, this mechanism is probably crucial for all forms of perceptual grouping (Geisler & Super, 2000), and hence may be wide spread within the cortex.

#### 6.5. Top-down factors

The model of contour grouping proposed here is relatively simple and ‘bottom-up,’ except for the final decision stage where the ‘observer’ is assumed to select the stimulus interval (in the forced choice task) based upon a high-level rule (the longest group consistent with the possible contour locations; see Fig. 7). The

success of our simple model in predicting contour grouping performance is a bit surprising because the human perceptual systems are widely regarded, in both the psychophysical and computational literature, as having an exceptional ability to find structure and regularity under conditions of high stimulus uncertainty. On the other hand, there must be ‘top-down’ effects of memory, attention and learning in contour grouping, at least under some circumstances. For example, it is presumably the influences of memory, attention and learning that are responsible for the appropriate grouping of impoverished images such as R.C. James’ ‘Dalmatian dog’.

One way that memory, attention and learning could influence contour grouping, within the present model, would be via modulation of the binding criterion  $\beta$  (Geisler & Super, 2000). In particular, the visual system could modulate (locally or globally) the binding criterion, and hence create many different patterns of grouping. The different groups formed during this process could be continuously matched against memory, and presumably, the visual system could select (or, with experience, could learn) those values of the binding criterion that produce strong matches with memory. Of course, this hypothesis is highly speculative, but it would be a relatively simple and powerful way for memory, attention and learning mechanisms to interact with low-level perceptual grouping mechanisms.

#### 6.6. Image tracing

We have shown that it is possible to measure some of the complex statistical regularities of the visual environment that form the basis for higher-order perceptual processes such as contour grouping. Of particular promise for future research is the image tracing method, which should enable investigators to measure the statistical regularities underlying other forms of perceptual grouping such those involving motion, binocular disparity, shading, and color. These statistical regularities should provide further fundamental insights into the brain mechanisms underlying perceptual grouping.

Another potential application of the image tracing methods (as well as the more direct methods) would be to analyze specific classes of visual environment. This might be particularly valuable for gaining insight into the visual systems of organisms that live in restricted environments. Although we found strong similarities in the statistics across the different visual images we analyzed, there were differences. Such differences may not be particularly important for understanding the design of visual systems that must cope with a wide range of environments, but they could be important for understanding visual systems that evolved within more constrained environments.

## Acknowledgements

We thank Robert Sekuler, Duane Albrecht, Lawrence Cormack, Steven Dakin and David Field for helpful comments. Supported by grant R01EY11747 from the National Eye Institute, National Institutes of Health, Bethesda MD.

## Appendix A

### A.1. Edge extraction

The transfer function used for finding the zero crossings was

$$H(u,v) = \exp\left(-\alpha\left(\frac{0.5 \log(u^2 + v^2) - \log(f_c)}{B_0}\right)^2\right), \quad (\text{A1})$$

where  $u$  and  $v$  are horizontal and vertical spatial frequency,  $f_c$  is the peak spatial frequency of the filter,  $B_0$  is the octave bandwidth of the filter, and  $\alpha$  is the constant  $-5.770780164$ . Each image was Fourier transformed, multiplied by this transfer function, and then inverse transformed. Pixels were examined one at a time, in a raster order. A pixel was regarded as a zero-crossing pixel if either the next pixel within the row, or the pixel immediately above the pixel in the next row, had a different sign.

The transfer function of the oriented filters was

$$H(u,v) = \exp\left(-\alpha\left(\frac{\log(|u'|) - \log(f_c)}{B_0}\right)^2 + i\phi \operatorname{sign}(u') - \log(2)\left(\frac{v'}{f_c \tan(0.5b_\theta)}\right)^2\right), \quad (\text{A2})$$

where  $u' = u \cos\theta + v \sin\theta$ ,  $v' = v \cos\theta - u \sin\theta$ ,  $\theta$  is the peak orientation,  $\phi$  is the spatial phase, and  $b_\theta$  is the orientation bandwidth. (Note that the ‘sign’ function has a value of  $+1$  for positive arguments and a value of  $-1$  for negative arguments.) For each orientation, two filtered images were obtained,  $I_\theta$  and  $\hat{I}_\theta$ , one for the filter in cosine phase ( $\phi = 0$ ) and one for the filter in sine phase ( $\phi = \pi/2$ ). These two filtered images were each squared and then summed to obtain the contrast energy, at that orientation, for each image pixel  $(x, y)$ :

$$E_\theta(x, y) = I_\theta(x, y)^2 + \hat{I}_\theta(x, y)^2. \quad (\text{A3})$$

Finally, the normalized energy response for each orientation, at each pixel, was obtained by dividing the contrast energy by the sum of the contrast energies for all orientations:

$$R_\theta(x, y) = \frac{E_\theta(x, y)}{E_{50} + \sum_\theta E_\theta(x, y)}. \quad (\text{A4})$$

The value of  $E_{50}$  was set so that the half saturation contrast, for a sine wave grating with the optimal frequency ( $f_c$ ), equaled the average half saturation contrast for neurons in primary visual cortex of monkey (40%). Similarly, the spatial frequency bandwidth and orientation bandwidth of the filters in Eq. (A2) were set so that the spatial frequency bandwidth and the orientation bandwidth of the response Eq. (A4) equaled the average values in visual cortex of monkey (1.5 octaves and  $40^\circ$ , respectively). A zero crossing pixel was considered ‘significant’ if the maximum response across orientation exceeded a criterion value of 10% of the maximum response to the optimal frequency  $f_c$ . The orientation at a significant edge pixel,  $\hat{\theta}$ , was obtained by finding the best fitting line to the response distribution across orientation, plotted on the unit circle:

$$\hat{\theta} = \tan^{-1}(S \pm \sqrt{1 + S^2}), \quad (\text{A5})$$

where,

$$S = \frac{\sum_\theta (R_\theta \sin \theta)^2 - \sum_\theta (R_\theta \cos \theta)^2}{\sum_\theta R_\theta^2 \sin \theta \cos \theta}. \quad (\text{A6})$$

### A.2. Histogram correction

To accurately estimate the co-occurrence probabilities, it is necessary to correct for the variations in bin size and for the effect of the finite circular aperture of the image. Specifically, the polar bin size increases with distance from the reference line. Further, the boundaries of the polar bins do not fall neatly on the square lattice of image pixels, producing some additional variation in bin size. Finally, the circular aperture of the image has an effect because when the distance between elements is greater than the distance from the element to the edge of the aperture, then the range of possible directions between the elements is restricted. Fortunately, it is possible to correct for all of these effects with the following simple procedure, which yields a correction factor that only depends upon the distance between elements. The procedure involves comparing every pixel (within the circular image aperture) with every other pixel (within the circular image aperture). Each of these comparisons consists of computing the Euclidean distance between the pixels, determining the distance bin,  $d$ , within which the distance falls, and then incrementing the counter for that bin,  $C(d) = C(d) + 1$ . The final counts for each distance bin serve as the correction factor. Thus, the absolute co-occurrence probability density function is given by

$$p(d, \theta, \phi) = \frac{N(d, \theta, \phi)1/C(d)}{\sum_d \sum_\theta \sum_\phi N(d, \theta, \phi)1/C(d)}, \quad (A7)$$

where  $N(d, \theta, \phi)$  is the number of edge element pairs falling in the 3D histogram bin  $(d, \theta, \phi)$ . Equivalent formulas hold for the conditional probability density functions.

### A.3. Random element display generation

Our method of generating contours and background regions is different from what is common in the literature, but it is more flexible. Here, the method is described in detail.

First, we pick a radius  $R$  for the whole display region, and an exclusion radius  $r$ , of arbitrary size, which defines a circular region centered on the element. Display elements are not allowed to be closer than a distance of  $2r$ . (Note that it is possible to allow element overlap by making  $r$  less than the element size, although this was not done in the present experiments.)

Second, the contour shape is specified mathematically in parametric form  $[x(s), y(s)]$ ,  $a \leq s \leq b$ . For example, the shapes of the filtered noise contours used here were first defined as modulations about the horizontal axis,

$$x_0(s) = s, \quad (A8)$$

$$y_0(s) = \sum_{i=1}^R T_i A_i \sin\left(2\pi \frac{i}{2R} s + \Phi_i\right), \quad (A9)$$

and then rotated by a random angle,

$$x(s) = s \cos \Theta + \sin \Theta \sum_{i=1}^R T_i A_i \sin\left(2\pi \frac{i}{2R} s + \Phi_i\right), \quad (A10)$$

$$y(s) = -s \sin \Theta + \cos \Theta \sum_{i=1}^R T_i A_i \sin\left(2\pi \frac{i}{2R} s + \Phi_i\right), \quad (A11)$$

where  $A_i$  and  $\Phi_i$  are the random amplitude and phase for the  $i$ th sine wave component,  $T_i$  is the value of the transfer function (used to filter the contour shape) at the spatial frequency  $i/2R$ , and  $\Theta$  is the random orientation of the contour. In these equations, the units of  $s$ ,  $R$ ,  $x$  and  $y$  are pixels.

Third, a set of potential contour element locations  $(s_1, \dots, s_k)$ , with a spacing of one pixel, is obtained between the starting and ending points ( $a$  and  $b$ ), using the following algorithm:

$$k = 1$$

$$s_k = a$$

while  $s_k \leq b$

$$s_{k+1} = s_k + \frac{1}{\sqrt{x'(s_k)^2 + y'(s_k)^2}}$$

$$k = k + 1$$

end while

$$K = k - 1$$

where  $K$  is the number of potential contour element locations, and the functions  $x'(s)$  and  $y'(s)$  are the derivatives of the functions in Eq. (A10) and Eq. (A11):

$$x'(s) = \cos \Theta + \sin \Theta \sum_{i=1}^R T_i A_i 2\pi \frac{i}{2R} \cos\left(2\pi \frac{i}{2R} s + \Phi_i\right), \quad (A12)$$

$$y'(s) = -\sin \Theta + \cos \Theta \sum_{i=1}^R T_i A_i 2\pi \frac{i}{2R} \cos\left(2\pi \frac{i}{2R} s + \Phi_i\right). \quad (A13)$$

Note that the ratio of these derivatives gives the slope of the contour at  $s$ , making it easy to manipulate the orientation of the elements with respect to the local orientation of the contour.

Fourth, we compute the approximate length of the contour  $L_c$ , the approximate area of the contour region  $A_c$ , and the approximate area of the background region  $A_b$ :

$$L_c = K + 1 - \sqrt{(x(b) - x(s_{K+1}))^2 + (y(b) - y(s_{K+1}))^2} \quad (A14)$$

$$A_c = L_c \cdot 2r + \pi r^2 \quad (A15)$$

$$A_b = \pi R^2 - A_c \quad (A16)$$

Finally, elements are sampled along the contour and then in the background such that the density (samples per unit area) is the same. Specifically, candidate contour element locations are randomly sampled with replacement from the set  $(s_1, \dots, s_K)$ , until the desired element density is obtained. Candidate background element locations are then randomly sampled with replacement from the whole display region until the desired element density is obtained. In either case, if a candidate element location overlaps with (is within  $2r$  of) any previously selected contour or background element location then it is not used, and another sample is taken. This procedure produces contours that are invisible unless the contour elements differ in some way from the background elements. The quality of the sampling in the display is checked subjectively either by making all the elements circles or all the elements dots. Also, to eliminate any possible residual information due to the sampling procedure, a ‘contour’ is generated in both temporal intervals of the forced choice display sequence, but in the target-absent interval the ‘contour’ and background elements are statistically identical.

With this display generation procedure it is possible to measure contour grouping performance for arbitrary contour shapes and element types.

## References

- Beck, J., Rosenfeld, A., & Ivry, R. (1989). Line segregation. *Spatial Vision*, 4, 75–101.
- Bosking, W. H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *Journal of Neuroscience*, 17, 2112–2127.
- Brunswick, E., & Kamiya, J. (1953). Ecological cue-validity of 'proximity' and other Gestalt factors. *American Journal of Psychology*, 66, 20–32.
- Buchsbaum, G., & Gottschalk, A. (1984). Chromaticity coordinates of frequency-limited functions. *Journal of the Optical Society of America A*, 1, 885–887.
- Burton, G. J., & Moorehead, I. R. (1987). Color and spatial structure in natural scenes. *Applied Optics*, 26, 157–170.
- Coppola, D. M., Purves, H. R., McCoy, A. N., & Purves, D. (1998). The distribution of oriented contours in the real world. *Proceedings of the National Academy of Sciences USA*, 95, 4002–4006.
- Dakin, S. C. (1997). The detection of structure in Glass patterns: psychophysics and computational models. *Vision Research*, 37, 2227–2246.
- Dakin, S. C., & Hess, R. F. (1998). Spatial-frequency tuning of visual contour integration. *Journal of the Optical Society of America A*, 15, 1486–1499.
- Dixon, R. E. (1978). Spectral distribution of Australian daylight. *Journal of the Optical Society of America*, 68, 437–450.
- Elder, H., & Goldberg, R. M. (1998). Inferential reliability of contour grouping cues in natural images, European Conference on Visual Perception (Oxford). *Perception*, 27(Supplement), 11.
- Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*. New York, Wiley & Sons, Inc.
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4, 2379–2394.
- Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local 'association field'. *Vision Research*, 33(2), 173–193.
- Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review*, 96, 267–314.
- Geisler, W. S., & Albrecht, D. G. (1997). Visual cortex neurons in monkeys and cats: detection, discrimination, and identification. *Visual Neuroscience*, 14, 897–919.
- Geisler, W. S., Thornton, T., Gallogly, D. P., & Perry, J. S. (2000). Image structure models of texture and contour visibility. *Proceedings of the NATO Workshop on Search and Target Acquisition, RTO-MP-45*, 15/1–15/8.
- Geisler, W. S., & Super, B. J. (2000). Perceptual organization of two-dimensional images. *Psychological Review*, 107, 677–708.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston, Houghton Mifflin.
- Gigus, Z., & Malik, J. (1991). *Detecting curvilinear structure in images* (CSD Technical Report No. 91/619). UC Berkeley.
- Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics*. New York: Krieger.
- Hess, R. F., & Dakin, S. C. (1999). Contour integration the peripheral field. *Vision Research*, 39, 947–959.
- Judd, D. B., MacAdam, D. L., & Wyszecki, G. W. (1964). Spectral distribution of typical daylight as a function of correlated color temperature. *Journal of the Optical Society of America*, 54, 1031–1040.
- Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (1999). Dynamics of spatial summation in primary visual cortex of alert monkeys. *Proceedings of the National Academy of Sciences*, 96, 12073–12078.
- Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, 23, 141–221.
- Kohonen, T. (1989). *Self-organization and associative memory* (3rd). Berlin: Springer.
- Kovacs, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Sciences*, 90, 7495–7497.
- Knill, D. C. and Richards, W. (eds). (1996) *Perception as Bayesian Inference*, Cambridge University Press.
- Maloney, L. T. (1986). Evaluation of linear models of surface reflectance with small numbers of parameters. *Journal of the Optical Society of America A*, 3, 1673–1683.
- McIlhagga, W. H., & Mullen, K. T. (1996). Contour integration with colour and luminance contrast. *Vision Research*, 36(9), 1265–1279.
- Miller, K., Keller, J. B., & Stryker, M. P. (1989). Ocular dominance column development: analysis and simulation. *Science*, 245, 605–615.
- Parent, P., & Zucker, S. (1989). Trace inference, curvature consistency and curve detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 823–839.
- Pettet, M. W., McKee, S. P., & Grzywacz, N. M. (1998). Constraints on long range interactions mediating contour detection. *Vision Research*, 38, 865–879.
- Sha'ashua, S., & Ullman, S. (1988). Structural saliency: The detection of globally salient structures using a locally connected network. Paper presented at the Proceedings of the Second International Conference on Computer Vision.
- Simoncelli (1997). *Statistical models of images: compression, restoration and synthesis*. 31st Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, IEEE Computer Society.
- Sirosh, J., & Miikkulainen, R. (1994). Cooperative self-organization of afferent and lateral connections in cortical maps. *Biological Cybernetics*, 71, 66–78.
- Switkes, E., Mayer, M. J., & Sloan, J. A. (1978). Spatial frequency analysis of the visual environment: anisotropy and the carpentered environment hypothesis. *Vision Research*, 18, 1393–1399.
- Walker, G. A., Ohzawa, I., & Freeman, R. D. (1999). Asymmetric suppression outside the classical receptive field of the visual cortex. *Journal of Neuroscience*, 19, 10536–10553.
- Wertheimer, M. (1958). Principles of perceptual organization. In D. C. Beardslee, & M. Wertheimer, *Readings in Perception*. Princeton: Van Nostrand.
- Yen, S.-C., & Finkel, L. H. (1998). Extraction of perceptually salient contours by striate cortical networks. *Vision Research*, 38, 719–741.