

Ecological statistics of Gestalt laws for the perceptual organization of contours

James H. Elder

Centre for Vision Research, York University, Toronto,
Canada



Richard M. Goldberg

User Centred Design Laboratory, IBM Canada Ltd.,
Toronto, Canada



Although numerous studies have measured the strength of visual grouping cues for controlled psychophysical stimuli, little is known about the statistical utility of these various cues for natural images. In this study, we conducted experiments in which human participants trace perceived contours in natural images. These contours are automatically mapped to sequences of discrete tangent elements detected in the image. By examining relational properties between pairs of successive tangents on these traced curves, and between randomly selected pairs of tangents, we are able to estimate the likelihood distributions required to construct an optimal Bayesian model for contour grouping. We employed this novel methodology to investigate the inferential power of three classical Gestalt cues for contour grouping: proximity, good continuation, and luminance similarity. The study yielded a number of important results: (1) these cues, when appropriately defined, are approximately uncorrelated, suggesting a simple factorial model for statistical inference; (2) moderate image-to-image variation of the statistics indicates the utility of general probabilistic models for perceptual organization; (3) these cues differ greatly in their inferential power, proximity being by far the most powerful; and (4) statistical modeling of the proximity cue indicates a scale-invariant power law in close agreement with prior psychophysics.

Keywords: perceptual organization, computational modeling, natural image statistics, image coding, contours, edges, proximity, good continuation, similarity

1. Introduction

Perceptual grouping is the problem of aggregating primitive image features that project from a common structure in the visual scene. Nearly 50 years ago, [Brunswik and Kamiya \(1953\)](#) suggested that the classical Gestalt principles of perceptual grouping should be quantitatively related to the statistics of the natural world and presented some informal data on the subject. Here we apply this idea to the perceptual problem of contour organization.

We define visual contour grouping as the problem of integrating the local luminance edges or oriented curve tangents that lie on a common luminance boundary in the image. An important property of contours is their one-dimensionality: a curve can be defined as a function of one parameter (e.g., arc length). This implies an ordering of points on the curve. Thus for the perceptual organization of contours, we must recover not just an aggregation but an ordered sequence of discrete edges or tangents. Without this ordering, it is not possible to define many useful higher-order properties (e.g., curvature, closure, concavity, and convexity).

We pose the problem of contour grouping as a problem of probabilistic inference: the goal of the computation is to compute highly probable sequences of

local curve elements. Properties of these local elements may serve as useful cues for deciding which elements should be grouped together, and in what order. Here we consider three such properties: proximity, good continuation, and similarity (in brightness and contrast). These are probabilistic cues: each can provide evidence for a certain grouping of elements, but none can provide a decision with 100% confidence. In order to understand how these cues may be used to optimize the accuracy of perceptual grouping decisions, we must quantitatively characterize the statistics of these cues in natural images. To estimate these statistics, we employ an advanced interactive software tool that allows human observers to rapidly trace the contours they perceive in natural images. These data can then be compared to past and future psychophysical studies to determine the degree to which the human perceptual organization system is tuned to the statistics of the natural world.

In summary, our contributions are:

- We develop a Bayesian model for the probabilistic combination of multiple grouping cues (proximity, good continuation, and luminance similarity) to determine local (pair-wise) groupings between contour elements.

- We demonstrate how the required probability distributions can be estimated from natural images using advanced software tools.
- We show that, when properly defined, these three cues are approximately independent.
- We formally characterize the inferential power of each of these three cues, and show that they differ dramatically in their importance for perceptual organization.
- We examine the variation in statistics over our sample of images to evaluate the utility of incorporating knowledge of these statistical models for general visual inference.
- We report a striking agreement between the statistics of the proximity cue in natural images and prior psychophysical results on the role of proximity in the perception of dot lattice stimuli (Oyama, 1961).
- We use these statistics to develop a parametric, generative model for natural image contours that can be used for both analysis and synthesis.

Our findings have important implications for human vision. There exists a wealth of psychophysical data for perceptual grouping phenomena. Qualitative and quantitative models have been constructed to account for these data. However, to answer the question of why these cues act and interact in a specific way requires a detailed understanding of the visual world to which our visual systems have been tuned by evolution and/or learning. Our hope is that this work will contribute to that understanding.

In the remainder of this section, we discuss previous psychophysical studies and computational models of contour grouping, and very recent attempts to characterize relevant statistics. In Sections 2-4, we develop the computational framework within which our study is conducted. In Sections 5 and 6, we report the methods and results for our study. Implications of these results are discussed in Sections 7 and 8.

1.1 Psychophysics of Contour Grouping

The early phenomenological demonstrations of the Gestalt psychologists led to the identification of distinct “principles” or cues to perceptual organization. There are several of these that apply to contour organization, including proximity, good continuation, and similarity (Wertheimer, 1923/1938; Koffka, 1935). Many of the Gestalt demonstrations showed that perceived organizations were determined by the cooperative or competitive interaction between these cues. However,

these demonstrations are qualitative, and it has been noted that to predict perceived organizations in novel displays, quantitative models of the relative strength of these cues are required (Hochberg, 1974; Kubovy & Holcombe, 1998).

1.1.1 Proximity

The cue of proximity is perhaps the most fundamental of the Gestalt grouping laws (Kubovy & Holcombe, 1998), and early quantitative studies suggest that the human visual system is extremely sensitive to proximity cues. Uttal, Bunnell and Corwin (1970) found the detection of dotted lines in random dot fields to depend strongly on the density of dots along the line. Barlow (1978) found psychophysical efficiencies of up to 50% for the detection of dot density changes in two-dimensional random dot displays.

These two studies raise the question of whether the law of proximity in the grouping of local elements sampling a one-dimensional contour can be considered as simply a limiting case of the action of proximity in two-dimensional texture grouping (Zucker, Stevens, & Sander, 1983). The fact that Barlow (1978) found no evidence for greater efficiency in detecting regions of higher dot density when these regions were elongated supports this idea.¹ On the other hand, more recent experiments by Elder and Zucker (1998a) demonstrate that the perceptual organization of fragmented figures is very sensitive to whether dots are placed on the boundary or in the interior of fragmented figures.

Oyama (1961) made what is probably the first attempt to quantify the law of proximity. Employing a regular rectangular dot lattice, he measured the proportion of time observers experienced vertical versus horizontal organizations as a function of the relative vertical and horizontal spacing. He found that the ratio of durations could be accurately modeled as a power law of the ratio of distances. The data also indicate a significant bias toward vertical organizations.

Kubovy and colleagues (Kubovy & Wagemans, 1995; Kubovy & Holcombe, 1998) have recently modified and elaborated this technique. Instead of a power law, they modeled their results using an exponential model of dot spacing, scaled by the minimum dot spacing present in the display.² They found evidence that the strength of the proximity cue increases with increasing stimulus duration (from 100 msec to 200 msec). They also found evidence for scale invariance in their experiments: equal scaling of the horizontal and vertical separations had little effect on their results.

Based on a number of experiments, Zucker and Davis (1988) have proposed that the perceptual organization of dotted contours changes abruptly at a dot:space ratio of roughly 1:5. They found that contours sampled more densely than this generate a number of classical illusions that sparsely sampled contours fail to generate. However,

no evidence for such a threshold is found in Kubovy's scaling results.

Elder and Zucker (1994) conducted a series of visual search experiments in which the target and distractors were fragmented outline shapes. They found that the psychophysical effects of the fragmentation could be characterized by the L_2 norm (sum of squares) of the gaps in the figures. This is consistent with a probabilistic model for contour grouping in which gaps are considered independent and the probability of grouping follows a half-Gaussian distribution. Elder and Zucker (1996a) later used this model in a computer vision algorithm for grouping closed contours in natural images. Gaussian distributions have also been used to model the proximity cue in the perception of clusters in two-dimensional random dot displays (Oeffelen & van Vos, 1982). Compton and Logan (1993) tested the Gaussian model for proximity against an exponential model, but could not statistically discriminate between them.

To summarize, the evidence suggests that proximity acts as a powerful, possibly scale-invariant cue for the perceptual organization of textures and contours. However, there is little agreement about the best quantitative description of the proximity cue: power law, exponential, and Gaussian models have been proposed and supported by various psychophysical data. An understanding of the statistics of the proximity cue for natural image contours might help toward focusing and eventually resolving this debate.

1.1.2 Good Continuation

Less is known about the quantitative nature of the Gestalt law of good continuation. Certainly we now have objective evidence for the action of this law. For example, Beck, Rosenfeld, and Ivry (1989) found longer reaction times for detection of a straight arrangement of elements in a random array when the line elements were laterally jittered. But the action of the good continuation cue in isolation from other cues has best been illustrated by the experiments of Field, Hayes, and Hess (1993). In these experiments, observers were asked to detect the presence of a curvilinear sequence of oriented elements in a random element field. Care was taken to equate the density of elements along and around the contour with the density in the whole display.³ Detection performance was found to decline for more wandering contours, and to rapidly descend to chance when the local elements themselves were jittered in their orientation relative to the path. These results suggest a powerful role for the cue of good continuation in the absence of a cue of proximity.

1.1.3 Brightness and Contrast Similarity

Even less is known about the quantitative role of brightness and contrast similarity in contour grouping. An early dot lattice experiment by Hochberg and Hardy

(1960) showed that proximity ratios of up to 2 can be overcome by intensity cues, and Earle (1999) has found evidence for a role of contrast similarity in the grouping of dots leading to the perception of Glass patterns. Contrast reversals between dot pairs are known to completely eliminate the perception of Glass patterns (Glass & Switkes, 1976).

1.2 Contour Grouping in Computer Vision

The computational problem of contour organization has been approached in many different ways. Cocircularity constraints have been applied within an iterative discrete relaxation framework to refine local curve estimates using global contour information (Zucker, Hummel, & Rosenfeld, 1977; Parent & Zucker, 1989). Energy minimization methods for approximating contours with spline models have been extensively investigated (Kass, Witkin, & Terzopoulos, 1987; David & Zucker, 1990). Multi-scale smoothness criteria have been used to impose an organization on image curves (Lowe, 1989; Saund, 1990; Dudek & Tsotsos, 1997), sequential methods for tracking contours within a Bayesian framework have been developed (Cox, Rehg, & Hingorani, 1993), and parallel methods for computing local "saliency" measures based on contour smoothness and total arc length have been studied (Sha'ashua & Ullman, 1988; Freeman, 1992; Alter, 1995; Williams & Jacobs, 1997).

While computational models for contour grouping generally exploit only geometric cues (proximity, good continuation), Elder and Zucker (1996a) have recently developed a probabilistic framework for contour grouping that also exploits luminance cues. In this work, maximum likelihood contours are estimated using a shortest path (Dijkstra's) algorithm. A similar approach has recently been taken by Crevier (1999), who extends the framework to allow grouping of circular arcs as well as straight contour segments, and attempts to relax the strict assumption of independence between the grouping of segment pairs.

In general, these techniques are capable of grouping edge points into extended chains. However, the goal of computing complete bounding contours has proven to be more elusive. Although approaches using global grouping cues such as convexity (Jacobs, 1996; Huttenlocher & Wayner, 1992) and closure (Elder & Zucker, 1996a; Mahamud, Thornber, & Williams, 1999) have yielded limited success, the general problem of computing the complete bounding contour of an object of arbitrary shape in a complex natural image remains essentially unsolved. One possible way of improving probabilistic models for contour grouping is to ground the models in the actual statistics of grouping cues for natural image contours.

1.3 Statistics of Contour Grouping

The connection between human visual processing and the statistics of the natural visual world has been recognized for nearly 50 years. [Attneave \(1954\)](#) and [Barlow \(1961\)](#) suggested that early visual processing may be designed primarily to reduce the statistical redundancy in natural images, thereby increasing coding efficiency. This idea has led to demonstrations that simple principles of redundancy reduction and sparseness of coding predict early visual filters that are qualitatively similar to the receptive field properties of early visual neurons ([Srinivisan, Laughlin, & Dubs, 1982](#); [Field, 1987](#); [Atick & Redlich, 1992](#); [Olshausen & Field, 1996](#); [Olshausen & Field, 1997](#)).⁴

The original proposals of [Attneave \(1954\)](#) and [Barlow \(1961\)](#) have led to considerable progress in our understanding of early visual processing in terms of linear transformations that reduce the redundancy or increase the sparseness or statistical independence of neural responses. However, there have been relatively few attempts to understand higher-level visual problems, such as perceptual organization, in terms of the statistics of the natural visual world. This is somewhat surprising because just prior to Attneave and Barlow's proposals, [Brunswik and Kamiya \(1953\)](#) proposed that the classical Gestalt principles of perceptual organization should be quantitatively related to the statistics of the natural world, and presented some data on the proximity of parallel contours in natural images. However, their suggestion remained largely untested until 1998, when [Kruger \(1998\)](#) first reported data on the second-order spatial statistics of Gabor filter responses to natural images, and we first reported the results of this study ([Elder & Goldberg, 1998a; 1998b](#)). More recently, there has been an interesting study of natural image statistics relevant to the problem of image segmentation ([Martin, Fowlkes, Tal, & Malik, 2001](#)), and two studies of natural image statistics relevant to the perceptual organization of contours ([Geisler, Perry, Super, & Gallogly, 2001](#); [Sigman, Cecchi, Gilbert, & Magnasco, 2001](#)). We discuss the studies relevant to contours in more detail below.

[Kruger \(1998\)](#) examined the second-order "co-occurrence" spatial statistics of Gabor filter responses to natural images. Filter responses were nonlinearly normalized and thresholded prior to statistical analysis. Kruger found statistical evidence for colinearity and parallelism relations in these second-order spatial statistics.

[Sigman et al. \(2001\)](#) also examined the co-occurrence statistics of oriented filter responses to natural images. They reported long-range correlations that adhered to the geometric principle of cocircularity. In a related study, [Geisler et al. \(2001\)](#) measured the co-occurrence statistics of oriented edge elements in natural images, and related these to human performance in detecting sampled contours in cluttered displays. They proposed a simple

model for grouping based in part on these statistics, and found that their model was to some degree consistent with the psychophysical data.

The correlations in the joint statistics of oriented edge elements observed in these studies reveal interesting parallel, colinear, and cocircular structure. Characterization of this statistical structure could be useful for understanding key aspects of early visual processing. Just as earlier statistical studies predicted early visual filters that are qualitatively similar to the receptive field properties of early visual neurons, these later studies may help us to relate natural image statistics to more complex aspects of neural coding, including the spatial nonlinearities in complex cells and lateral interactions between neurons in early visual cortex ([Gilbert & Wiesel, 1989](#)). An understanding of this second-order statistical structure is also useful for image processing applications, including image compression and image denoising ([Simoncelli, 1997](#)).

However, we will argue that these statistics are not sufficient to understand the statistical basis for the perceptual organization of contours. The heart of the matter is that perceptual grouping is not simply the problem of detecting correlations. Rather, the problem is to integrate the sequences of elements that project from common structures in the scene. In particular, contour grouping is the problem of integrating those edge elements that lie on a common luminance boundary. There are other reasons one might predict statistical correlations between edge elements: between parallel elements in texture flows, for example, or between colinear elements on different components of a regular texture. Because the statistics reported in these studies result from a mixture of these effects, they cannot be used directly to understand contour grouping per se.

In our study, we use human observers to trace the contours in natural images, and thus obtain pairs of contour elements that we know should be directly grouped. At the same time, we randomly sample the image to obtain pairs of tangent elements that should not be grouped. From the perspective of probabilistic inference, it is vital to have statistics for both (contour and random) events: it is the ratio of the likelihood distributions for these two events that determines the posterior probability for contour grouping ([Section 4](#)). Since these two events are not distinguished in the co-occurrence statistics collected in other studies, these statistics are insufficient for the probabilistic inference of contours.

In addition to collecting co-occurrence statistics, [Geisler et al. \(2001\)](#) used a form of contour tracing in order to distinguish the statistics relating contour elements on a common contour from those relating elements on different contours. (Our study was first reported [[Elder & Goldberg, 1998a; 1998b](#)] two years before the first report of the work of Geisler et al. [[Geisler, Super, & Gallogly, 2000](#)]). However, their

technique differs from ours in one crucial respect. Their traces indicate which elements are perceived to lie on a common contour, but they do not provide any information about the ordering of the elements along the contour. This is important because a defining property of contours is their one-dimensionality: a contour may be parameterized by a single real variable, e.g., $\alpha(s) = (x(s), y(s))$. This imposes an ordering on the local elements of a curve. For example, the point $\alpha(b)$ on the curve lies between points $\alpha(a)$ and $\alpha(c)$ if and only if $a < b < c$. These properties are essential for defining higher order properties of curves, e.g., curvature, closure, concavity, and convexity.

To maintain this one-dimensional characteristic in a discrete encoding, a contour must be represented as an ordered sequence of local elements. In the study of Geisler et al., contours are represented not as ordered sequences but as unordered sets of oriented elements, and their statistics relate arbitrary pairs of tangents on a contour. These statistics therefore do not reflect the fundamental one-dimensional topological property of contours.

In contrast, we define the problem of contour grouping as the recovery of sequences of tangents projecting from the contours of a scene. Participants trace sequences of tangents defining the contours they perceive in natural images. From these data, we can derive statistics for tangent pairs that are successive components of a common contour. These statistics thus inform us about the cues to inferring the sequence of tangents defining perceived contours.

We model these sequences as Markov chains. The Markov approximation captures the local nature of the physical processes that give rise to these contours, and is consistent with the monotonically decreasing nature of the autocorrelation function of natural images, the spatiotopic structure of early visual cortex, and the well-studied psychophysical principle of proximity. Application of the Markov approximation allows us to understand the problem of contour grouping by characterizing the statistics of local grouping between successive elements comprising a contour.

Our model reflects the fact that the statistical dependencies between neighboring tangents on a contour are much stronger than those between distant tangents on a contour. (Here the terms “neighboring” and “distant” refer to ordinal distance in the chain, not Cartesian distance in the image.) In the approach of Geisler et al. (2001), the power of the strong statistics relating neighboring tangents is diluted with the weak statistics relating distant tangents. This leads to substantial differences between our statistics and their statistics, as we shall see.

It should be noted that there are potential disadvantages of contour-tracing methods. Compared with the measurement of ordinary second-order (co-occurrence) statistics, tracing methods introduce

additional possible sources of error and bias. This may include biases of the participants doing the tracing, and errors caused by the software that manages the tracing process. A nice aspect of the Geisler et al. study is that they use both methods, and compare the results of the two. Of course, because we expect significant differences even without error, it is not possible to verify either method with this comparison. Our view is that we have no real choice but to use human-traced contours, because it is not possible to get the required statistics for the Bayesian inference of contours without ground-truth data, and human traces are the best available approximation to ground truth data.

Another distinction of our study is the multiplicity of grouping cues explored, and the rigorous manner in which they are compared. It has long been understood that perceptual organization is determined by the simultaneous action of several factors or principles (Wertheimer, 1923/1938; Koffka, 1935). How are these different factors combined? What is the relative importance of these factors in determining the perceived organization? Other studies have investigated two grouping factors (proximity and good continuation) but have not quantified the relative importance or independence of each individually, and photometric cues have been completely ignored.

Here we investigate three of the classical Gestalt principles (proximity, good continuation, and luminance similarity) for the organization of local curve elements into extended contours. We investigate these properties separately so that we may estimate their relative inferential power, but we also study to what degree they provide independent information for contour grouping, and how they can be optimally combined.

2. Local Contour Representation

To measure the statistics of contour grouping cues in natural images, we must first be able to detect and represent the local contour elements. In other studies, edges have been detected using fixed-scale filters followed by simple nonlinearities. For example, Kruger (1998) used fixed-scale oriented Gabor filters, followed by a point nonlinearity and thresholding. Sigman et al. (2001) used fixed-scale steerable quadrature-pair filters to measure the local oriented energy, followed by a threshold. Geisler et al. (2001) used a two-stage filtering process. In a first stage, potential edge locations were identified as the zero-crossings in the response of a nonoriented log Gabor function. The local energy at these locations was then measured using oriented quadrature-pair log Gabor filters, and a threshold was applied.

Although the filters used in these edge detection techniques all bear some resemblance to the receptive fields of early visual neurons, they are likely to be gross

oversimplifications of the cortical processing involved in edge detection. Neurons in primary visual cortex are extremely diverse in their receptive field properties, even within a single class of cell. For example, foveal simple cells in primary visual cortex of macaque range in peak spatial frequency from roughly 0.5 cycles per degree (cpd) to more than 16 cpd, spatial frequency bandwidth from roughly 0.4 octaves to more than 2.6 octaves, orientation bandwidth from less than 10 deg to more than 180 deg, and receptive field height:width ratios from roughly 1:1 to 16:1 (DeValois, Albrecht, & Thorell, 1982; Parker & Hawken, 1988).

Many psychophysical studies using adaptation, masking, and subthreshold summation techniques have demonstrated that early visual processing results from the activity of multiple mechanisms with different spatial frequency tunings (Campbell & Robson, 1968; Wilson & Bergen, 1979; Watson, 1982; Watt & Morgan, 1984; Wilson & Gelb, 1984; Watson, 2000). Recent work suggests that psychophysical edge detection also requires mechanisms over a broad range of orientation bandwidths, as is suggested by the physiological data (Sachs & Elder, 2000).

In this work, we use a multi-scale edge detection method developed for computer vision applications (Elder & Zucker, 1998b). In some respects, this method is more biologically plausible than the methods used by Kruger (1998), Sigman et al. (2001), and Geisler et al. Most critically, filters varying over a range of spatial frequencies (scales) are employed, similar to the range found in early visual cortex of primate. The adaptive filter selection method has been found to predict human visual acuity for blurred edges (Elder & Zucker, 1996b) and human detection efficiency for windowed edges in noise (Sachs & Elder, 2000).

However, our goal here is not to propose a model for human visual edge detection, nor to prefilter the images through a simple model of early visual cortex. Rather, our goal is to reliably detect and locally represent the contours projecting from luminance transitions in the scene. Only if this is achieved will we be accurate in our estimates of contour statistics in natural images. Further, if we hope to eventually measure the degree to which the human visual system is tuned to the statistics of natural images, it is vital that we do not corrupt our measurements of natural images with biases induced by our simplistic models of cortical processing. Such a circular procedure would undermine the significance of any links we may discover.

What leads us to choose the multi-scale edge detection algorithm of Elder and Zucker (1998b) is thus not its biological plausibility, but its performance in detecting edges over the broad range of blur, contrast, and clutter found in natural images. The multi-scale nature of the algorithm is crucial to achieving this.

One important advantage of this representation is that we can invert it to reconstruct an approximation of the original image, and thus can subjectively and

objectively measure any information lost or distortions introduced (Elder, 1999). We have conducted detailed studies to show that this representation is perceptually nearly complete, with minimal loss of information.

Although our local edge computation yields an accurate representation of the image, these edges do not provide an optimal basis for contour grouping. Due primarily to spatial discretization, the edge map representation of a contour is jagged and noisy. In order to avoid these problems, we employ a second level of representation in which the contour is locally represented by tangents with position, orientation, and length represented by real numbers (Elder & Zucker, 1996a). These two stages are described in more detail below.

2.1 Edge Computation

We model local edges as Gaussian-blurred step discontinuities in image intensity (Figure 1). The model consists of 5 parameters (Elder & Zucker, 1998b):

- Location (to the nearest pixel in this implementation)
- Orientation
- Blur scale σ_b
- Asymptotic intensity on the bright side of the edge I_+
- Asymptotic intensity on the dark side of the edge I_-

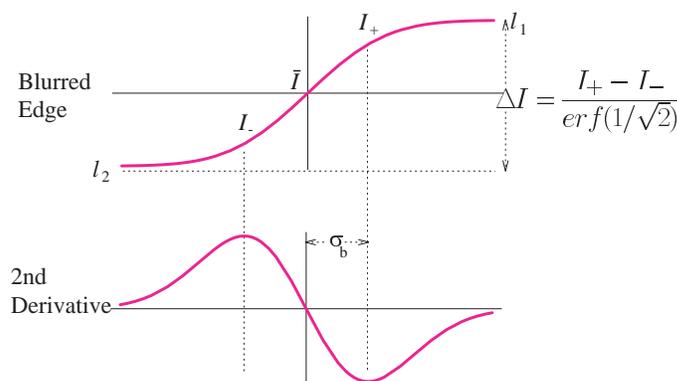


Figure 1. Local Gaussian-blurred edge model and derivative-based detection/estimation mechanisms.

Detection of edges and estimation of model parameters are based on measurement of the gradient of the intensity function using steerable first derivative of Gaussian filters (Freeman & Adelson, 1991; Perona, 1995), and on estimation of the locations of zero-crossings and extrema of the second derivative using steerable second derivative of Gaussian filters, steered in the gradient direction. While the zero-crossing of the second derivative localizes the edge, the separation of the 2nd derivative extrema in the gradient direction is used to estimate the blur scale of the edge. Estimates of the image intensity at the 2nd derivative extrema are used to

estimate the mean intensity $\bar{I} = (I_+ + I_-)/2$ and the magnitude of the intensity change $\Delta I = (I_+ - I_-)/(\text{erf}(1/\sqrt{2}))$ at the edge. These are then used to estimate the asymptotic intensities $I_1 = \bar{I} + \Delta I/2$ and $I_2 = \bar{I} - \Delta I/2$ on either side of the edge (Figure 1).

A major obstacle to reliable edge detection is the scale problem: how to choose the scale of local estimation filters in order to prevent false positives and distortion due to noise, while minimizing distortion caused by neighboring image structure. Our method for edge detection solves this problem with an adaptive scale space technique called local scale control (Elder & Zucker, 1998b). This technique selects, at each point in the image, the minimum reliable scale for local estimation. At this scale, hypotheses concerning the sign of response of a linear filter at each point can be tested with statistical reliability. This means in turn that zero-crossings can be reliably detected and localized. This theory of scale selection has been shown to accurately predict human psychophysical performance in edge localization and blur

estimation tasks (Elder & Zucker, 1996b). An example of the edge map produced by this algorithm is shown in Figure 2b.

Because our interest is to estimate the properties of grouping cues for the actual contours in an image, it is important that our local elements (edges) be accurate. Otherwise we may simply be measuring artifact of our edge detection methodology. Because we lack ground truth for the actual location of contours in the image, a direct estimate of the accuracy of our edges is not available. We must therefore consider indirect methods.

We have recently reported a method for inverting our edge representation to compute an estimate of the original image from which the edge map was computed (Elder, 1999). Using this algorithm, we have shown our edge representation to be both objectively and subjectively accurate for a wide variety of images. Figure 2c shows the reconstruction of the image in Figure 2a from the computed edge representation.

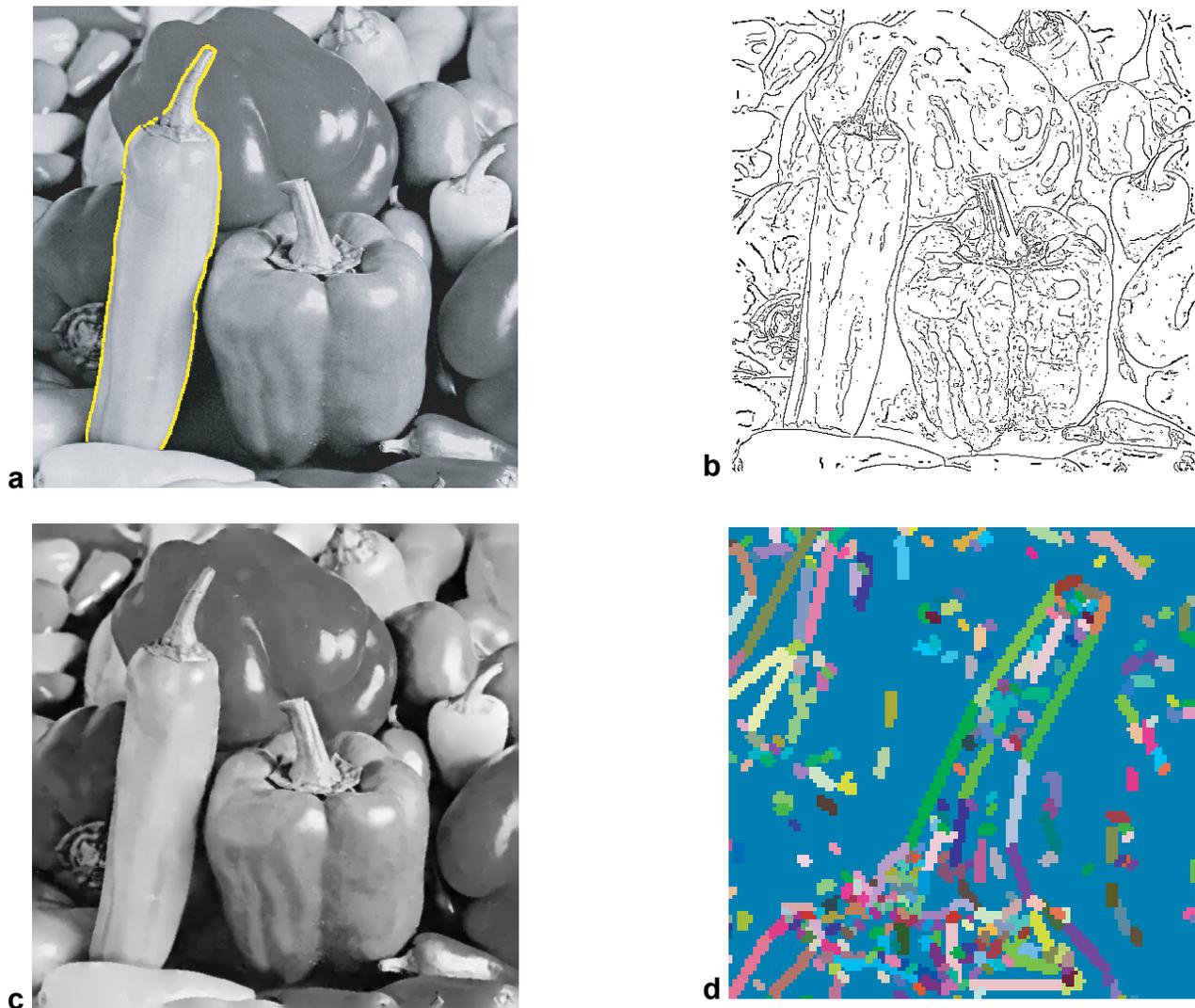


Figure 2. a. Example contour traced by a human participant. b. Edge map from which contours are defined. c. Reconstruction of image from edge representation. d. Portion of tangent map. Each colored line segment represents a distinct tangent.

Although our local edge computation yields an accurate representation of the image, these edges do not provide an optimal basis for contour grouping. The problem is illustrated in Figure 3. Due primarily to spatial discretization, the edge map representation of a contour is jagged and noisy. Even when a set of edges are known to be generated by the same contour, it is difficult to specify an appropriate ordering on the edge pixels, and tracing a contour through any particular ordering yields a curve corrupted by high curvature wiggles due to the discretization.

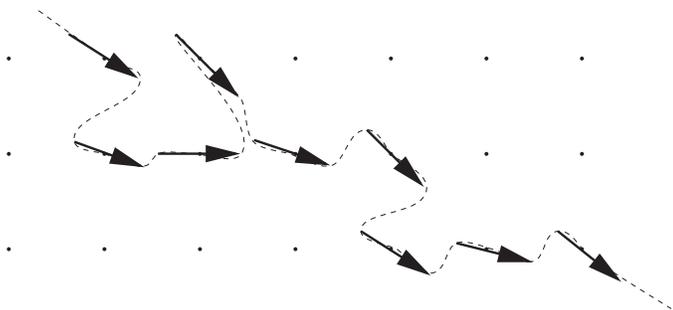


Figure 3. Observable geometric cues between edge pixel tangents are dominated by artifact introduced by the spatial discretization of the image.

2.2 Tangent Representation

In order to avoid these problems, we employ a second level of representation in which the contour is locally represented by tangents with position, orientation, and length represented by real numbers (Elder & Zucker, 1996a). These tangents, not constrained by the discrete pixel grid, and often averaging over multiple, roughly colinear edges, provide a much more accurate basis for one-dimensional perceptual grouping. We stress that these computations are not intended as a model for biological visual processing. Rather, they are intended simply to provide an accurate estimate of local contour information.

To construct the tangent representation, each local edge in the image generates a tangent line passing through the edge pixel in the estimated tangent direction. The tangent estimates that are 8-connected to the local edge, which lie within an ε -neighbourhood of the local tangent line, and whose gradient directions are compatible with that of the local edge, are identified with the extended tangent model. For this study, we use $\varepsilon = 1.5$ pixels. Gradient direction compatibility is determined based on the known level of sensor noise, using a first-order noise propagation model.

A greedy algorithm is used to select the subset of tangents that will represent the image contours. Given a connected set of local edges, the longest line segment that faithfully models a subset of these is determined. This subset is then subtracted from the original set. This

process is repeated for the connected subsets thus created until all local edges have been modeled. Luminance estimates for the edge pixels modeled by each tangent are averaged, and each tangent is thus represented as a 6-element vector:⁵

- x position
- y position
- Length
- Orientation
- Luminance on light side of tangent
- Luminance on dark side of tangent

By convention, the spatial component (first four elements) of each tangent vector is represented as a 90 deg counterclockwise rotation from the gradient direction. The (x,y) position represents the location of the base of the vector in the image. A portion of the tangent map computed for the image in Figure 2a is shown in Figure 2d.

3. A Probabilistic Model for Tangent Grouping

The set of tangents T computed from an image may be enumerated: $T = \{t_1, \dots, t_N\}$. The set S of possible contours may then be represented as tangent sequences: A sequence of tangents $s = \{t_{\alpha_1}, \dots, t_{\alpha_n}\} \in S$ if and only if

$$\alpha_i \in \{1, \dots, N\}, \forall i \in \{1, \dots, n\},$$

$$\alpha_i \neq \alpha_j, \forall i \in \{1, \dots, n\}, j \in \{2, \dots, n-1\}, i \neq j. \quad (1)$$

This definition restricts the mapping to be injective (tangents cannot be repeated in the sequence), with the exception that the first and last tangent may be the same. In this case, the contour is closed. For the purposes of this study, we will restrict our attention to contours for which contrast polarity does not reverse along the contour, thus tangents in a contour are linked “tip-to-tail.”

We assume that there exists a correct organization of the image $C \subset S$. Correctness may be defined in terms of objective ground truth, e.g., the contours that bound the objects in a scene. Unfortunately, except perhaps for highly simplified artificial or synthetic scenes, objective ground truth is difficult to obtain. Because our interest is in the perceptual organization of typical natural images, we elect in this study to define correctness in terms of human perception, i.e., a contour is correct if it is what a human observer perceives. If this aspect of human perception is close to veridical, then our study reveals aspects of how contours in the natural world appear in

images. If not, we can at least say that our measurements reveal aspects of the information likely used by the human visual system to group contours.

A visual system may use a number of observable properties D to decide on the correctness of a hypothesized contour. Here we examine properties corresponding to the classical Gestalt cues of proximity, good continuation, and similarity. Knowing these properties D influences the probability $p(c \in C | D)$ that a particular contour c is correct.

Here we are interested in how properties $d_{ij} \in D$ defined on pairs of sequential tangents $\{t_i, t_j\}$ may influence the probability that a contour $c = \{t_{\alpha_1}, \dots, t_{\alpha_n}\}$ is correct. The local property $d_{ij} \in D$ may, e.g., represent the distance between the two tangents or a measure of the curvature of the best continuant between the two tangents. Note that these local properties do not embody many important aspects of global geometry and topology. In general, the visual system may also apply one or more global constraints that only a subset of contours may satisfy, e.g., closure, simplicity (no self-intersections), and completeness. However, here we focus only on characterizing the statistics of local cues for grouping.

Using Bayes' theorem, the probability $p(c \in C | D)$ that a particular contour c is correct may be written as

$$p(c \in C | D) = \frac{p(D | c \in C)p(c \in C)}{p(D)} = \frac{1}{1 + (LP)^{-1}} \quad (2)$$

where

$$L = \frac{p(D | c \in C)}{p(D | c \notin C)} \quad \text{and} \quad P = \frac{p(c \in C)}{p(c \notin C)}. \quad (3)$$

In general, the prior ratio reflects the expected number of tangents in the contours of the image and can be modeled fairly easily. However, because contours can be many tangents in length, the likelihood distributions are in general of very high dimension; in order to model the statistics of contour grouping, some simplifying approximations must be made. Here we model contours as Markov chains (Mumford, 1992; Elder & Zucker, 1996a; Williams & Jacobs, 1997), so that tangent grouping is pair-wise independent. In particular, we will assume that only the grouping cues $d_{ij} \in D$ directly relating tangents on the hypothesized contour c depend upon the hypothesis, and that these are conditionally independent. Then

$$L = \prod_{i=1}^n \frac{p(d_{\alpha_i \alpha_{i+1}} | c \in C)}{p(d_{\alpha_i \alpha_{i+1}} | c \notin C)} \quad (4)$$

and the likelihood ratio can be computed as a product of local likelihood ratios. Note that this model makes no assumption that local tangent groupings are unique.

The intuition behind this Markov approximation is that the strongest statistics lie in the relations between directly successive tangents on the contour, so these

should be modeled directly. The weaker statistics relating more distant tangents are captured approximately through the Markov structure.

Depending on the nature of the global constraints, it may be possible to compute maximally probable contours using an efficient shortest-path computation on a directed graph representing the Markov network. In prior work (Elder & Zucker, 1996a), we developed an algorithm for computing closed contours using these approximations. While in the present work we are interested principally in the problem of estimating local probability distributions, we will employ interactive grouping software that makes use of these approximations in order to rapidly infer contour segments between tangents selected by human observers (Section 5).

4. Defining the Cues

Here we focus on local statistics, and so in the following we will consider contours consisting of just two tangents $\{t_i, t_j\}$. Then we have

$$p(\{t_i, t_j\} \in C | D) = \frac{1}{1 + (LP)^{-1}} \quad (5)$$

where

$$L = \frac{p(d_{ij} | \{t_i, t_j\} \in C)}{p(d_{ij} | \{t_i, t_j\} \notin C)}, \quad P = \frac{p(\{t_i, t_j\} \in C)}{p(\{t_i, t_j\} \notin C)}. \quad (6)$$

The prior ratio P is approximately equal to the probability that two arbitrarily selected tangents are grouped. If we assume that pair-wise groupings are typically (but not always) unique, then P is approximately equal to the reciprocal of the number of tangents in the image.

The likelihood ratio L represents the ratio of the likelihood of the observables given that t_i and t_j are directly grouped to the likelihood given that they are not. We will refer to these likelihoods throughout the paper as the *contour* and *random* likelihoods, respectively.

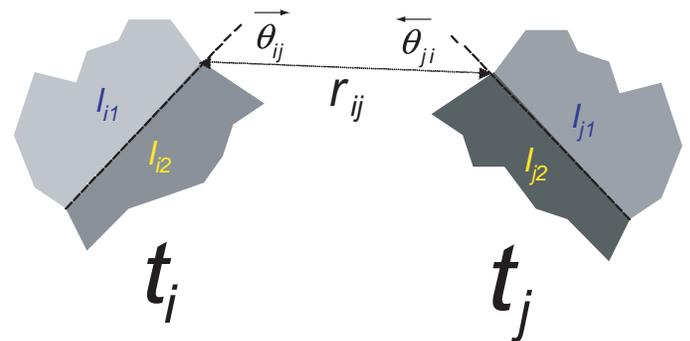


Figure 4. Observable data relating two tangents. See text for details.

In this study we consider three observable cues that we expect to be most influential on the probability of grouping: proximity, good continuation, and similarity. As a first order approximation, we use a rectilinear model of completion between two tangents t_i, t_j (Figure 4):

- *Proximity*: A function of the length r_{ij} of the straight-line interpolant (gap).
- *Good continuation*: A function of the two orientation changes $\bar{\theta}_{ij}$ and $\bar{\theta}_{ji}$ induced by the interpolation.
- *Similarity*: A function of the differences in estimated image intensities l_{i1}, l_{i2} and l_{j1}, l_{j2} between the two tangents. In this study, we consider only grouping that preserves the contrast polarity of the contour.

If we can approximate distinct cues (proximity, good continuation, etc..) as independent when conditioned upon grouping hypotheses, the contour and random likelihoods can be factored.⁶ Given m distinct cues $d_{ij}^k, k \in 1..m$ relating tangents t_i and t_j , we then have:

$$p(d_{ij} | \{t_i, t_j\} \in C) = \prod_{k=1}^m p(d_{ij}^k | \{t_i, t_j\} \in C) \quad (7)$$

$$p(d_{ij} | \{t_i, t_j\} \notin C) = \prod_{k=1}^m p(d_{ij}^k | \{t_i, t_j\} \notin C).$$

It is these likelihoods that we wish to estimate in the present study.

5. Methods

5.1 Participants

Five unpaid participants, all undergraduate or graduate students of vision science, participated in the experiment. All had normal or corrected-to-normal vision. The participants were aware of the goals of the study.

5.2 Apparatus

Experiments were conducted on a Pentium workstation with a Sony Trinitron display. Proprietary software, discussed in detail in Section 5.4, was employed to display the images and allow participants to trace perceived contours.

5.3 Stimuli

Nine arbitrarily selected natural grayscale images were employed (Figure 5). An attempt was made to include images of diverse subjects and settings (e.g., people, objects, animals; indoor, outdoor).

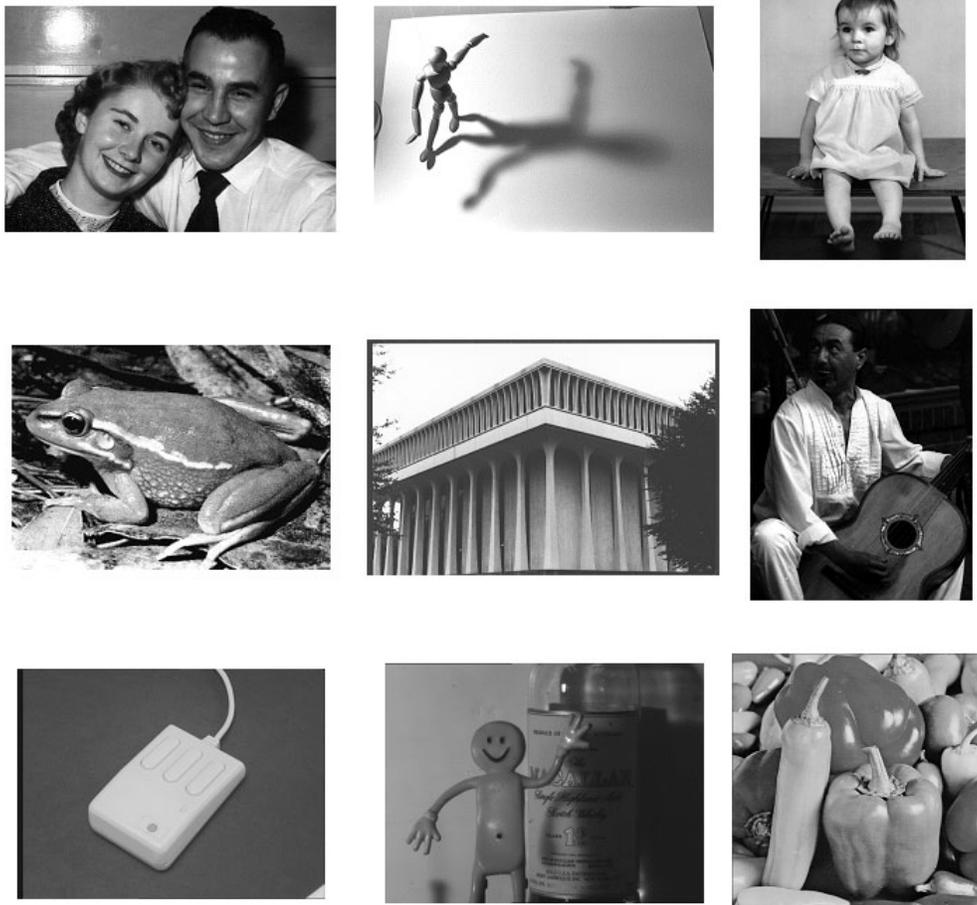


Figure 5. Images used for our experiments.

5.4 Software Tool for Interactive Contour Grouping

Our goal was to estimate the probability distributions for the observable grouping cues available in the contours perceived by human observers. To do this, we needed a method for translating observer percepts into tangent sequences: somehow observers must be able to trace the contours they see, and each trace must be mapped to a sequence of tangents.

In order to allow participants to accurately and efficiently trace contours, we employed a software package called Interactive Contour Editor (ICE), previously developed for a demonstration of contour-based image editing technology (Elder & Goldberg, 2001). ICE represents an image by information at its edges, and then allows the image to be modified by direct editing of the contours. This technology uses previously developed algorithms for reconstructing images from our edge representation (Elder, 1999). In order to allow users to efficiently manipulate contours, ICE provides an interactive contour grouping mechanism based on the tangent representation and independence approximations described in Sections 3 and 4 (Elder & Zucker, 1996a). The likelihood distributions employed are generally Gaussian, with parameters chosen using a combination of common sense and trial and error.⁷

Rather than requiring experimental participants to painstakingly trace each tangent of a contour in sequence, we used the grouping feature of ICE as a kind of “power-assist” to accelerate the process. This approach has a number of advantages:

- Accurate estimation of probability distributions requires a large amount of data. Using ICE, participants can group a long sequence of tangents with a relatively small number of mouse clicks, allowing the required quantity of data to be collected quickly.
- The increase in efficiency reduces observer fatigue, and thus may improve the quality of data.
- ICE turns approximately positioned mouse clicks into selections of the nearest tangent, and allows the observer to group contours in chunks. These capabilities eliminate the need for zooming and unzooming the image, which can cause the observer to lose global perspective and can introduce errors into the data.

A potential disadvantage of the methodology is that ICE may itself introduce errors into the data by selecting groupings that are not perceived by the observer. This problem was largely avoided by provision of an “undo” mechanism that allowed participants to delete groupings they had not intended to make. Participants were

instructed to use the ICE grouping tool to full advantage, but to be constantly vigilant for such errors and to correct them immediately. Typically the errors made by ICE were “blunders” that were difficult to miss.

The graphical user interface (GUI) for ICE is shown in Figure 6c. Both the working image and edge map are displayed. For this study, we made use only of the features of ICE that allow contour tangents in a natural image to be selected and grouped as a sequence.

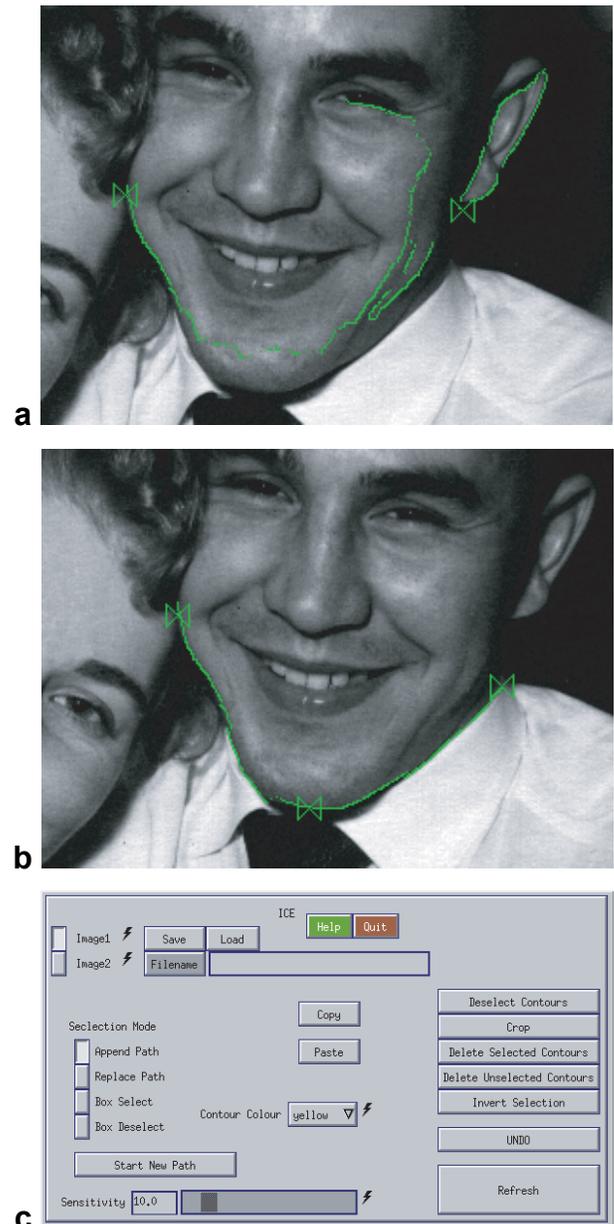


Figure 6. a. Example of incorrect path computed by ICE. b. The correct path is computed when the participant selects more closely spaced tangents. c. Graphical user interface for ICE.

Participants select contours by clicking on either the image or the edge map. Grouping is initiated by clicking near a contour. This click initiates a nearest neighbor

search in the area of the mouse click to find the nearest edge point. The coordinates of the nearest edge point are used to index the tangent map and thus obtain the index of the tangent corresponding to the edge point. The selected tangent is highlighted in color on both edge and image displays. When the user clicks near a second edge point, a terminating tangent index is similarly obtained.

These two tangent indices form input to a graph algorithm that determines the most probable sequence of tangents connecting the two selected tangents, under the independence approximations discussed in Sections 3 and 4 (Elder & Zucker, 1996a).

In append path mode, subsequent mouse selections will append maximum likelihood contour segments to the previously computed path. In replace path mode, a third mouse selection will deselect the previous path and begin a new path at the selected edge point.

Figure 6 shows an example of this interactive grouping procedure. Selected tangents are indicated by bow tie markers. Because the grouping algorithm is imperfect, selecting two tangents that are too distant may lead to a nonsense path (Figure 6a). In such cases, the participant may undo the path and, with ICE in append path mode, select a sequence of more closely spaced points along the contour that the algorithm can more easily connect (Figure 6b).

5.5 Procedure

Each participant was instructed to use the ICE software to trace all of the contours they perceived in each of the natural images. Images were presented in a random order. Participants could select tangents by clicking in either the image or the edge map. Participants were instructed to try to group complete contours, but not to group multiple contours together. Participants were also instructed to consider not only the contours bounding objects, but also contours arising from reflectance changes, shading and shadows.

Unlike Geisler et al. (2001), we did not force the participants to trace all automatically detected contours in the image. Thus the potential exists that participants may have traced only the more salient contours, even though they were instructed to trace all contours they perceived, and this may lead to bias in the statistics. The difficulty in forcing the participants to trace all detected contours is that, depending upon the characteristics of the monitor, there may be some contours they simply cannot see due to blur, low contrast, and clutter. Geisler et al. get around this by imposing an arbitrary response threshold on edge detection filters and thus suppressing low-contrast edges, but, of course, this could also introduce bias in the contours that are traced.

The experiment produced a total of 16,222 tangent pairs perceived to be directly grouped. Many of these tangent pairs were selected by more than one participant: thus only 7,476 of these pairs were unique. We

considered the set of tangent pair samples selected by each observer to be an independent random sample from a common underlying population, and therefore used the full set of 16,222 pairs in estimating the contour distributions.

In order to estimate the random likelihood distributions, we randomly sampled 10,000 pairs of tangents from the same set of 9 images. No attempt was made to avoid tangent pairs that are perceived to be grouped, because such pairs form an insignificant proportion of the total number of tangent pairs present in an image.⁸

5.6 Modeling of Distributions

As a first-order approximation, we will assume that the grouping cues are mutually independent in the contour and random conditions, i.e., when conditioned upon $\{t_i, t_j\} \in C$ or $\{t_i, t_j\} \notin C$. Thus we are interested in modeling the individual marginal distributions for each of the cues.

We wish to estimate the likelihood distributions for each of the Gestalt cues d_{ij}^k , given tangents that are successive elements of the same contour ($p(d_{ij}^k | \{t_i, t_j\} \in C)$), and for tangents that are not ($p(d_{ij}^k | \{t_i, t_j\} \notin C)$). In this way we hope to quantify our understanding of the classical Gestalt laws. For example, intuitively we expect that the distance between two tangents known to be successive elements of the same underlying contour will tend to be smaller than the distance between random tangents, but we are seeking a more complete quantitative description of this intuition in the form of these two likelihood distributions.

6. Results

6.1 Proximity Cue

Figure 7a shows a log-log plot of the contour likelihood distribution $p(r_{ij} | \{t_i, t_j\} \in C)$, where r_{ij} is the separation between tangents. A scatterplot of the empirical distribution is shown in blue. Note that while we would expect the true likelihood distribution to decrease monotonically as a function of the distance between tangents, the data are nonmonotonic, peaking at roughly $r_{ij} = 2$ pixels. We believe the falloff observed for small gaps is because of small random errors in our algorithm's localization of tangent endpoints; we discuss this below.

For gaps greater than 2 pixels, the data appear roughly linear in log-log coordinates. In other words, the contour likelihood distribution for the proximity cue follows a power law:

$$p(x) = ax^{-b} \tag{8}$$

$$\text{where } x \in [x_0, \infty), \quad a = (b-1)x_0^{b-1}, \quad x_0 > 0, \quad b > 1.$$

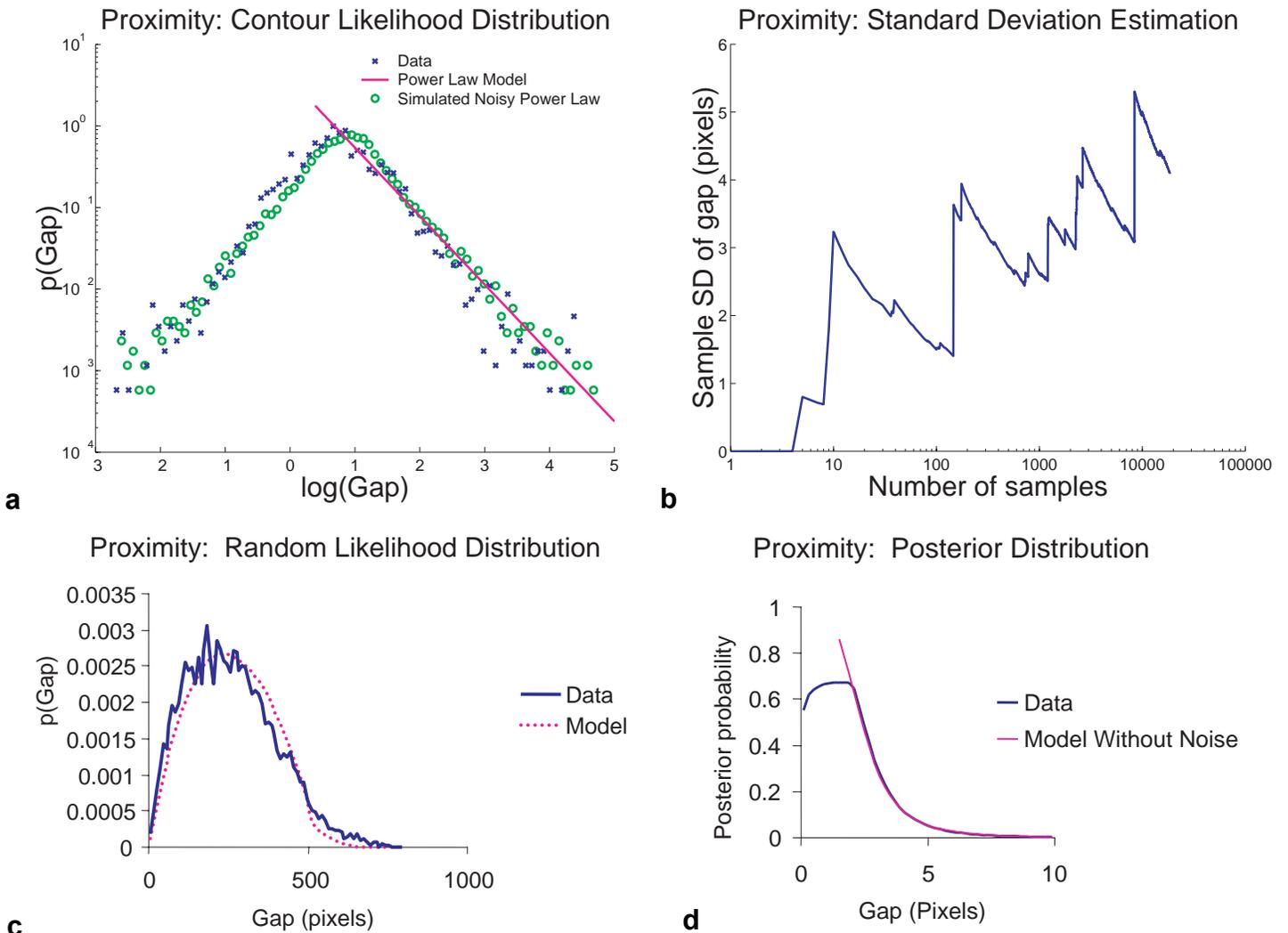


Figure 7. a. Estimated likelihood distribution $p(r_{ij} | \{t_i, t_j\} \in C)$ for the proximity cue between two tangents known to be successive components of a common contour. b. Sample standard deviation of tangent separation along contours, as a function of sample size. The lack of convergence suggests that the standard deviation is undefined. c. Estimated likelihood distribution $p(r_{ij} | \{t_i, t_j\} \notin C)$ for the proximity cue between two tangents selected randomly from the image. d. Practical (with estimation noise) and theoretical (without estimation noise) estimates of the posterior distribution $p(\{t_i, t_j\} \in C | r_{ij})$ for tangent grouping based on the proximity cue.

A maximum likelihood estimate of the underlying power law is shown in magenta in Figure 7a.

Bootstrapping to estimate standard errors, we estimate the power law parameters to be:

$$a = 3.67 \pm 0.09$$

$$b = 2.92 \pm 0.02$$

$$x_0 = 1.402 \pm 0.007 \text{ pixels}$$

Thus the gaps along a contour follow a power law, with a minimum distance between tangent endpoints of 1.4 pixels (roughly the distance between diagonally adjacent pixels), and an exponent of 2.92. While the mean separation is 2.9 pixels, the standard deviation and higher order moments are undefined: for example, the

model predicts that the sample standard deviation of the distance between tangents along a contour, estimated from observed data, will increase unbounded as a function of the size of the sample. Figure 7b shows that our data do indeed exhibit this behavior.

To model the complete distribution of the proximity law, including small gaps, we assume that this power law is corrupted by noise caused by small errors in localizing tangent endpoints. We have observed that these localization errors can be as large as ± 1 pixel in both horizontal and vertical directions. Modeling these random errors as independent and uniformly distributed, we can generate samples of the resulting noisy power law. Such a sample is shown in green in Figure 7a: the striking

similarity between the real and simulated data provides strong support for the model.

The random likelihood distribution $p(r_{ij} | \{t_i, t_j\} \notin C)$ for the proximity cue is shown in Figure 7c. To model this distribution, we assumed that tangents are uniformly distributed over the image. We then computed the exact distribution based on this assumption for a 100×100 pixel image. Distributions for square images of different sizes are obtained by simply scaling this distribution. Although our images were of various sizes and generally were not square, for the purposes of this study we

approximated the images as 512×512 pixels. The resulting model is seen to fit the data well.

Having models of the likelihood distributions for both contour and random conditions, and knowing the number of tangents in each image, we can use Equation 1 to compute the posterior probability $p(\{t_i, t_j\} \in C | r_{ij})$ as a function of the tangent separation. Figure 7d shows the posterior with and without the noise introduced by the tangent computation. It can be seen that for small separations, the grouping probability is very high.

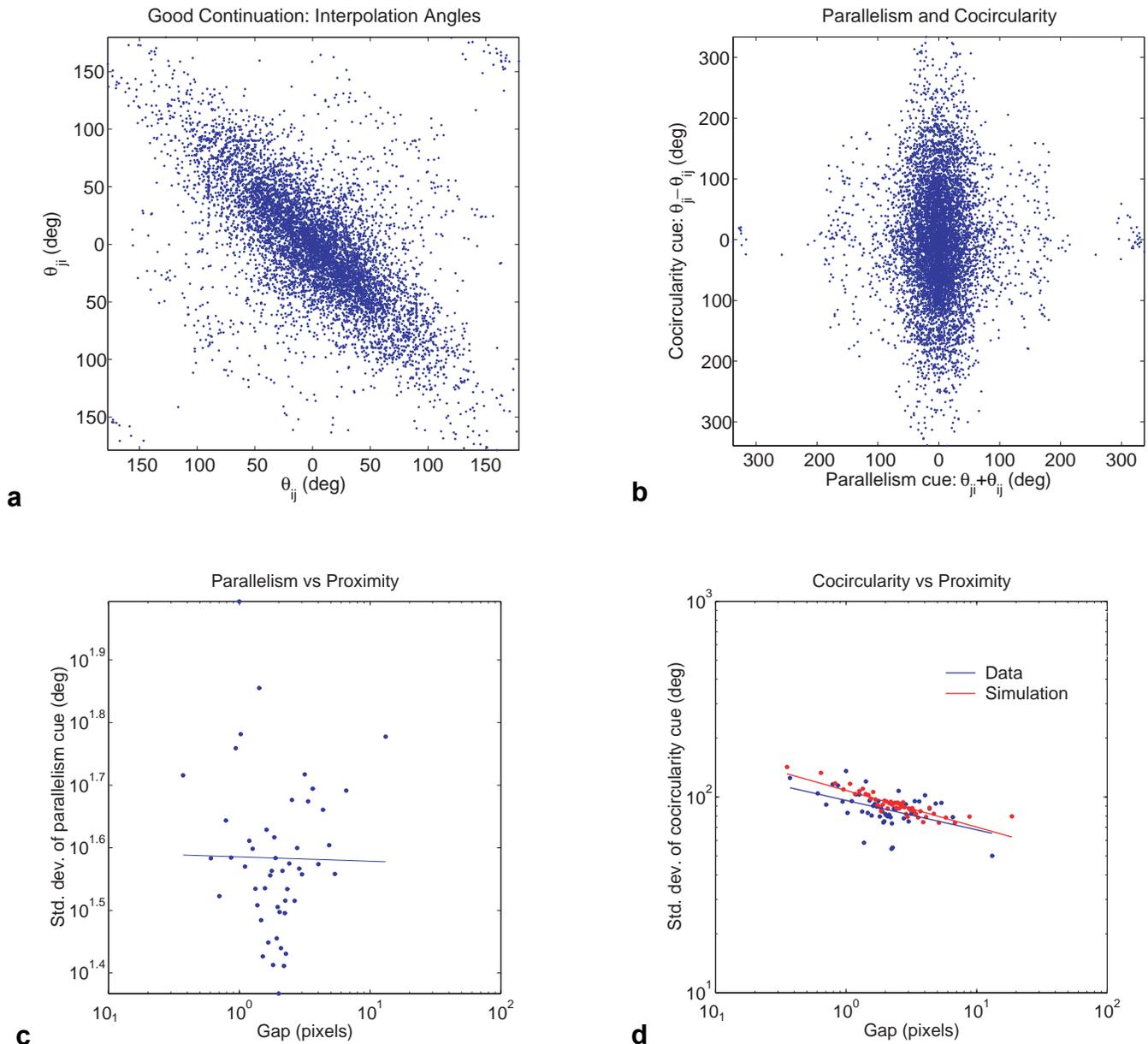


Figure 8. Coding the good continuation cue. All data are drawn from the contour condition. a. Scatterplot showing negative correlation of the two interpolation angles. b. Linear recoding into parallelism and cocircularity cues results in a more independent code. c. The parallelism cue is approximately uncorrelated with the proximity cue. d. The cocircularity cue is correlated with the proximity cue. See text for details.

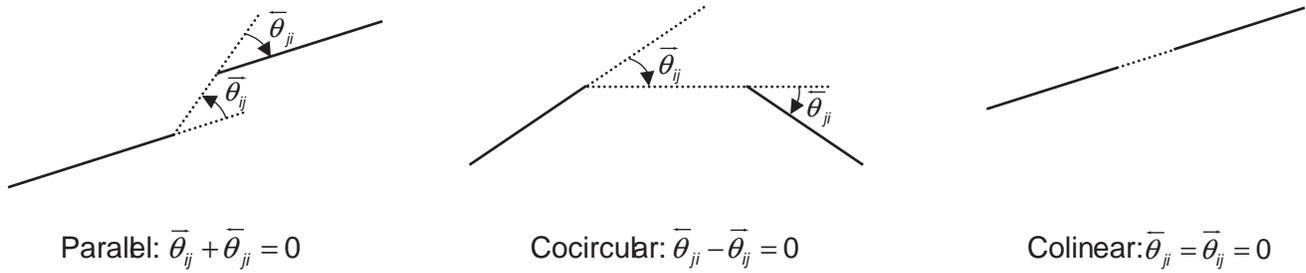


Figure 9. Recoding the interpolation angles into a sum and difference leads to an intuitive representation of good continuation.

6.2 Good Continuation Cue

Using our first-order model of contour continuation (Figure 4), the grouping of two tangents generates two interpolation angles $\bar{\theta}_{ij}, \bar{\theta}_{ji}$. We find that these two angles are strongly anticorrelated in the contour condition (Figure 8a). This suggests a recoding of the angles into sum ($\bar{\theta}_{ji} + \bar{\theta}_{ij}$) and difference ($\bar{\theta}_{ji} - \bar{\theta}_{ij}$) cues. This encoding appears to be close to the principal component basis for the good continuation cue: the new variables are approximately uncorrelated in the contour condition (Figure 8b).

There are four advantages to this new representation of the good continuation cue. First, it appears to be close to the principal components of the data in the contour condition. Second, when less than 180 deg in absolute value, these two new variables have very natural meaning, in terms of intuition and in terms of the literature (Figure 9). The sum variable represents parallelism: the two tangents are parallel if and only if $\bar{\theta}_{ji} + \bar{\theta}_{ij} = 0$, and $\bar{\theta}_{ji} + \bar{\theta}_{ij}$ increases monotonically in absolute value as the tangents become less parallel. The difference variable represents cocircularity: the two tangents are cocircular if and only if $\bar{\theta}_{ji} - \bar{\theta}_{ij} = 0$, and $\bar{\theta}_{ji} - \bar{\theta}_{ij}$ increases monotonically in absolute value as the tangents become less cocircular. Finally, note that two tangents are colinear if and only if both $\bar{\theta}_{ji} + \bar{\theta}_{ij} = 0$ and $\bar{\theta}_{ji} - \bar{\theta}_{ij} = 0$. Because for roughly 94% of our data for the contour condition both variables are less than 180 deg in absolute value, we will generally refer to $\bar{\theta}_{ji} + \bar{\theta}_{ij}$ and $\bar{\theta}_{ji} - \bar{\theta}_{ij}$ as parallelism and cocircularity cues, respectively.

In the 6% of cases where either $\bar{\theta}_{ji} + \bar{\theta}_{ij}$ or $\bar{\theta}_{ji} - \bar{\theta}_{ij}$ is greater than 180 deg in absolute value, we cannot think of them as measuring parallelism or cocircularity. This is because neither geometric property embodies the sense in which the contour is being traversed. Clearly this is an important constraint in contour grouping, and the $\bar{\theta}_{ji} + \bar{\theta}_{ij}$ and $\bar{\theta}_{ji} - \bar{\theta}_{ij}$ variables do take this into account.

This representation of good continuation is different from that used by Geisler et al. (2001). In their representation, while one angle represents parallelism, the other angle has no obvious intuitive meaning, and we find in our own data that these two angles are highly correlated.

A third advantage of the parallelism and cocircularity encoding of the good continuation cue is that sources of error in measuring these two variables are quite different, and it useful to separate these out. We discuss this at length below. The final advantage is that the parallelism cue will turn out to be much stronger in inferential power than the cocircularity cue (Section 7), and this supports our general goal of constructing representations that concentrate the greatest predictive power in the smallest number of variables.

The parallelism cue (as represented by its standard deviation) is also very nearly uncorrelated with the proximity cue (Figure 8c). It is thus appropriate to consider the marginal statistics of the parallelism cue. Figure 10a shows the likelihood distribution for the parallelism cue in the contour condition (blue curve). The distribution is kurtotic (kurtosis = 16.9). To model this distribution, we employed a generalized Laplacian distribution that has been used in the past to model kurtotic wavelet response histograms (Mallat, 1989; Simoncelli & Adelson, 1996; Simoncelli, 1999):⁹

$$p(x) = Ae^{-(c|x|/\sigma)^\gamma}$$

$$\text{where } c = \sqrt{\frac{\Gamma(3/\gamma)}{\Gamma(1/\gamma)}}$$

$$\text{and } A = \frac{c\gamma}{2\sigma^\gamma \Gamma(1/\gamma)}.$$
(9)

This distribution is symmetric and unimodal. σ is the standard deviation and $\gamma \in (0, \infty)$ controls the kurtosis. If $\gamma = 2$ the distribution is Gaussian. If $\gamma < 2$ the distribution has positive kurtosis. If $\gamma > 2$, the distribution has negative kurtosis, approaching a uniform distribution as $\gamma \rightarrow \infty$. To model an empirical distribution, we determine the generalized Laplacian distribution with matching standard deviation and kurtosis. Given a target kurtosis, the required γ is found using standard nonlinear optimization techniques.

The generalized Laplacian model for the parallelism cue is shown in red in Figure 10a. The model parameters for the parallelism cue in the contour condition are listed in Table 1.

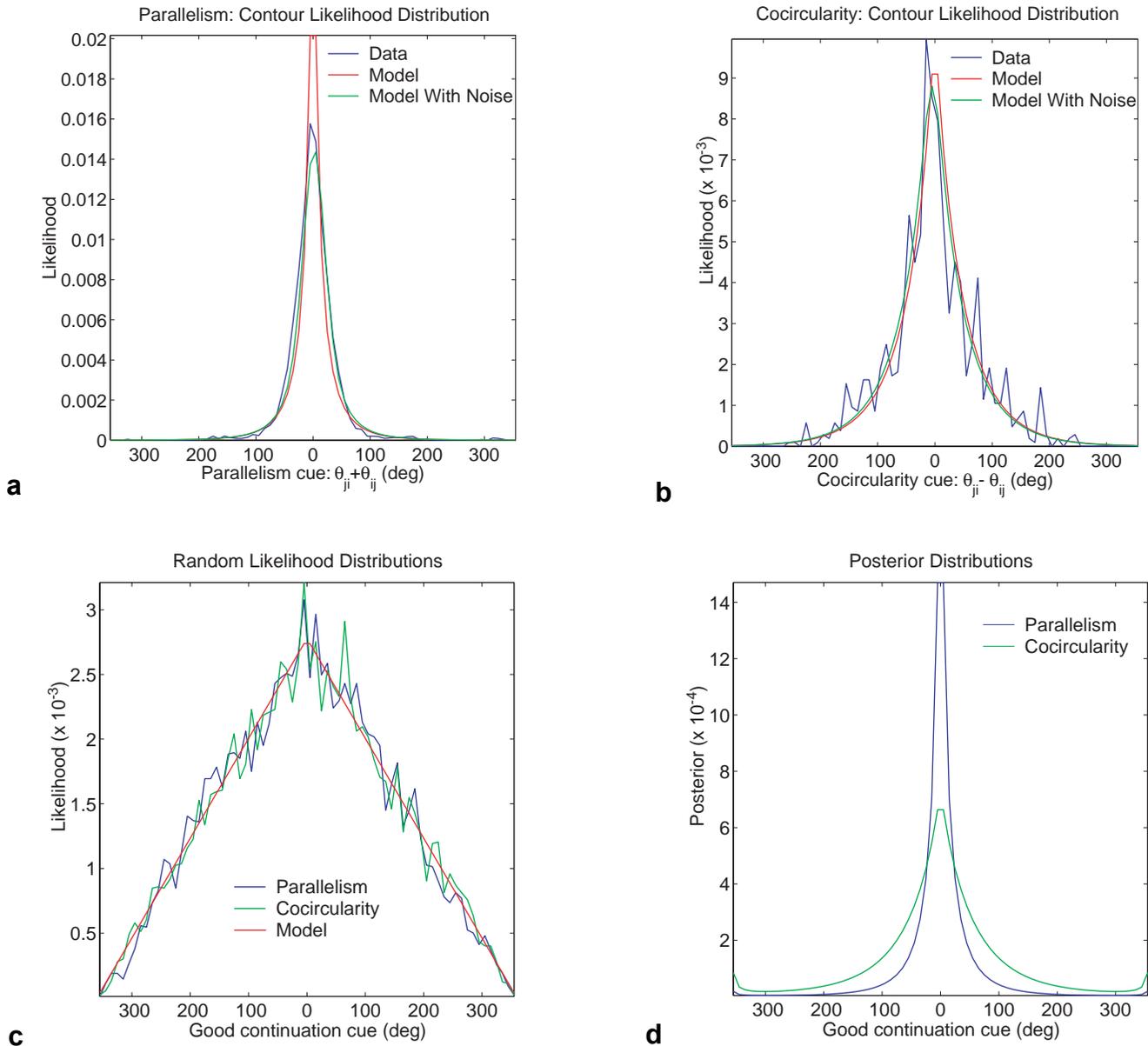


Figure 10. Statistical distributions for the good continuation cues. a. Likelihood distribution $p(\bar{\theta}_{ji} + \bar{\theta}_{ij} | \{t_i, t_j\} \in C)$ for the parallelism cue in the contour condition. b. Likelihood distribution $p(\bar{\theta}_{ji} - \bar{\theta}_{ij} | \{t_i, t_j\} \in C)$ for the cocircularity cue in the contour condition. c. Likelihood distribution for the good continuation cues $p(\bar{\theta}_{ji} + \bar{\theta}_{ij} | \{t_i, t_j\} \notin C)$ and $p(\bar{\theta}_{ji} + \bar{\theta}_{ij} | \{t_i, t_j\} \in C)$ in the random condition. d. Posterior distributions $p(\{t_i, t_j\} \in C | \bar{\theta}_{ji} + \bar{\theta}_{ij})$ and $p(\{t_i, t_j\} \in C | \bar{\theta}_{ji} - \bar{\theta}_{ij})$ for the good continuation cues. See text for details.

Table 1. Generalized Laplacian Parameters for Parallelism and Cocircularity Cues in the Contour Condition

	σ	Kurtosis	γ
Parallelism	42.1 deg	16.9	0.54
Cocircularity	76.8 deg	3.86	0.91

Although the parallelism cue is approximately uncorrelated with the proximity cue, this is not the case for the cocircularity cue (Figure 8d): the standard deviation of the cocircularity cue decreases as the distance

between tangents increases. In other words, the cocircularity cue is weaker for more proximal tangents.

We suspected that this observation stems from measurement error. While the parallelism cue depends only on the difference in estimated orientation of the two tangents, the cocircularity cue depends on the orientation $\bar{\phi}_{ij}$ of the virtual line connecting the relevant endpoints of the tangents (Figure 11). Denoting horizontal and vertical separations of the two tangents as r_x and r_y respectively, we have $\tan \bar{\phi}_{ij} = r_y / r_x$, and the partial derivatives of $\bar{\phi}_{ij}$ with respect to r_x and r_y are

$$\frac{\partial \bar{\phi}_{ij}}{\partial r_x} = -\frac{r_x}{r_x^2 + r_y^2} \quad \frac{\partial \bar{\phi}_{ij}}{\partial r_y} = \frac{r_y}{r_x^2 + r_y^2} \quad (10)$$

Thus estimation of the cocircularity cue is ill conditioned when the separation between tangents is small.

To minimize this source of error, we restricted our analysis of the cocircularity cue in the contour condition for tangent separations of 5 pixels or greater, where this small-separation effect is negligible. The data and generalized Laplacian model are shown in Figure 10b. The parameters of the model are listed in Table 1.

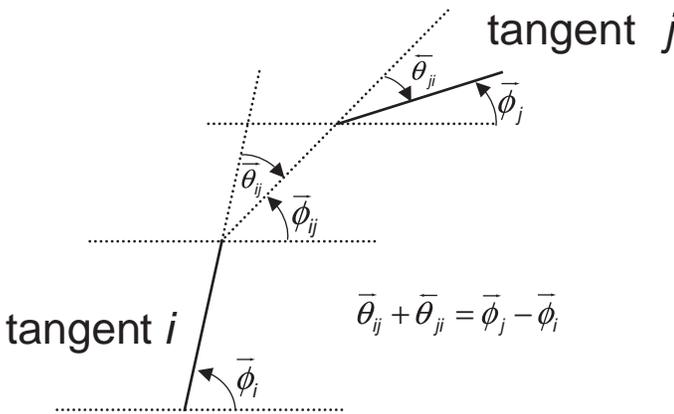


Figure 11. The parallelism cue (sum of interpolation angles) can be reexpressed as the difference in orientation of the two tangents. The cocircularity cue, however, depends on the orientation $\bar{\phi}_{ij}$ of the straight-line interpolant between the tangents.

The generalized Laplacian model can be seen to overestimate the likelihood for nearly parallel tangents (Figure 10a). We believe that a more accurate model may be obtained by modeling the noise in tangent orientation estimation. The effect of independent additive Gaussian noise of standard deviation σ_θ in the two tangent angles is to blur the likelihood distribution with a Gaussian blur kernel of scale $\sqrt{2}\sigma_\theta$. We therefore estimate the standard deviation σ_θ of the measurement noise by minimizing the least-squared difference between the data and Gaussian-blurred model for the parallelism and cocircularity cues, obtaining an estimate of $\sigma_\theta = 9.9 \text{ deg}$. The resulting models are shown in green in Figure 10a and 10b.

To determine whether localization error in tangent endpoints could account for the observed correlation between proximity and cocircularity cues, we used our model for endpoint error (uniform noise of ± 1 pixel in x and y coordinates, Section 6.1) and our model for the cocircularity cue for tangent separations greater than 5 pixels to simulate cocircularity data for a range of tangent separations. The result, shown in red in Figure 10d, is

quite consistent with the observed data, suggesting that were localization error eliminated, the cocircularity cue would be roughly uncorrelated with the proximity cue.

The likelihood distributions for the good continuation cues in the random condition can be modeled by assuming an isotropic tangent distribution. The resulting model can be seen to fit the data well (Figure 10c). Figure 10d shows the posterior distributions for the two cues. In deriving these distributions, we have attempted to remove errors in estimating tangent orientation and location. These distributions thus represent a “best case” scenario. Note that the parallelism cue appears to be more informative than the cocircularity cue: this will be studied more formally in Section 7.

6.3 Similarity Cue

The tangent representation includes an estimate of the image intensity on either side of each tangent. Differences in these intensities between tangents form a potential cue for contour grouping.

Tangents may be grouped in two ways, so that the polarity of contrast is either preserved or reversed along the contour. We did permit contrast reversals in the contours traced by our participants, and found that roughly 13% of the local groupings involved a contrast reversal. However, on examination it appeared that a number of these contrast reversals were erroneous. The difficulty was that our tracing software did not clearly indicate a contrast reversal to participants, who therefore had no way to detect and correct erroneous reversals. We therefore decided to restrict our analysis to segments of contour where no reversals were indicated.

One obvious way of encoding intensity similarity information is to consider the difference $l_{j1} - l_{i1}$ in the intensity of the light sides of the two tangents t_i, t_j as one cue, and the difference $l_{j2} - l_{i2}$ in the intensity of the dark sides of the two tangents as a second cue.

The problem with this approach is that these two cues are highly correlated in the random condition (Figure 12a) and therefore their joint distribution cannot be accurately approximated by the product of the marginal distributions. By inspection it appears that the first principal component of this joint distribution is roughly the sum of these two differences. This forms a brightness cue $b_{ij} = \bar{l}_j - \bar{l}_i = (l_{j1} + l_{j2} - l_{i1} - l_{i2})/2$, measuring the difference between the two tangents t_i, t_j in the mean luminance of the dark and light sides of the underlying edge. The second principal component then forms a contrast cue $c_{ij} = \Delta l_j - \Delta l_i = |l_{j1} - l_{j2}| - |l_{i1} - l_{i2}|$, measuring the difference in the amplitudes of the intensity steps at the two tangents. Using this new basis will result in approximate decorrelation of the cues for the nongrouped case (Figure 12b).

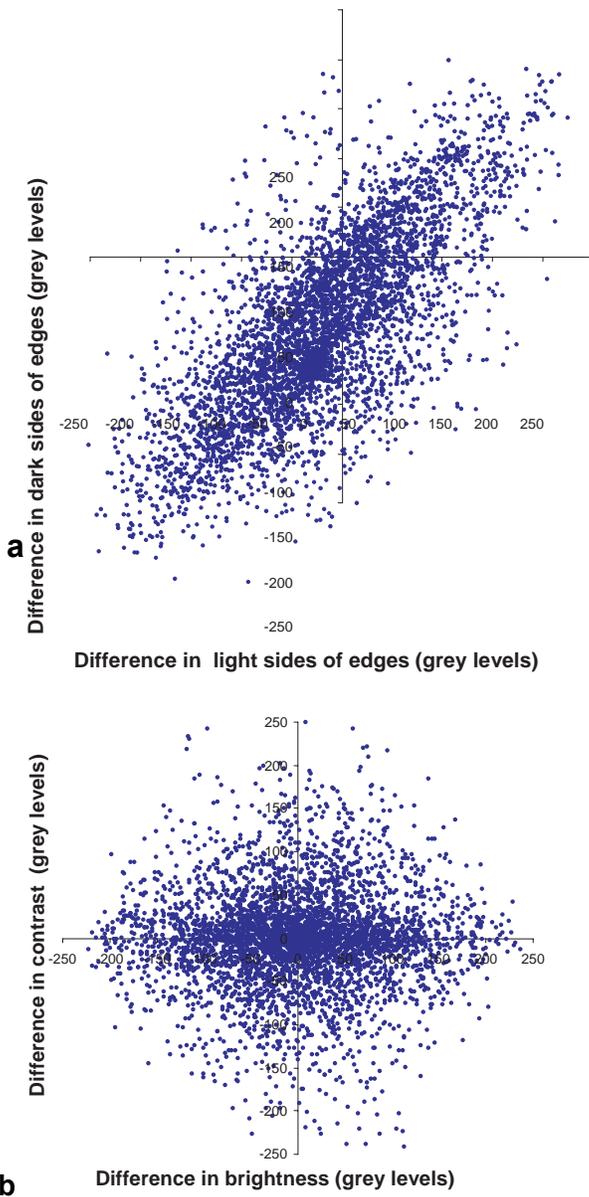


Figure 12. a. Dark and light luminance difference cues between randomly selected tangents are strongly correlated. b. Brightness and contrast cues for randomly selected tangents are approximately uncorrelated.

Table 2 lists the Pearson correlations for these various luminance cues in both the grouped and random conditions. Overall, the brightness and contrast cues are less correlated than the light/dark difference cues. However, we must be careful not to assume that these low correlations mean that the cues are independent. Table 3 lists the Pearson correlations for the absolute values of these same cues. Note the high correlation between brightness and contrast cues in the contour condition. Clearly, it would be a mistake to conclude that these two cues are independent.

These results present us with a dilemma. While the dark/light representation is superior for the contour condition, the brightness/contrast representation is superior for the random condition. One solution is to use different representations for the two conditions; however, then it would be impossible to quantify the inferential

power of the individual cues: the most we could do is quantify the power of the two luminance cues taken together.

Table 2. Comparison of Pearson Correlation Coefficients for Two Measures of Luminance Similarity

	Dark/light	Brightness/contrast
Contour	0.12	-0.06
Random	0.77	0.01

Table 3. Comparison of Pearson Correlation Coefficients for Absolute Values of Two Measures of Luminance Similarity

	Dark/light	Brightness/contrast
Contour	0.19	0.53
Random	0.65	0.02

Because one of the prime purposes of this study is to quantify the inferential power of contour grouping cues, we elect instead to use the brightness/contrast representation. As we shall see, the brightness cue turns out to be a far more powerful cue than the contrast cue, and so this decomposition allows the reduction of the luminance information into a single cue without substantial loss of inferential power.

We found that the contrast of our images varied considerably (standard deviations from 48 to 77 grey levels), and that the statistics of luminance grouping cues co-vary with image contrast (Pearson correlations of 0.72 for the brightness cue and 0.34 for the contrast cue: Figure 13). In order to increase the reliability, and therefore the inferential power of the luminance cues, it is useful to normalize by overall image contrast.

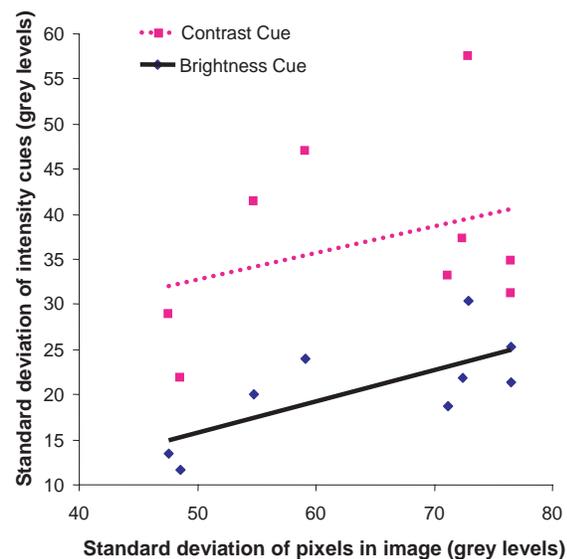


Figure 13. The statistics of both brightness and contrast cues are strongly correlated with the pixel statistics of the image.

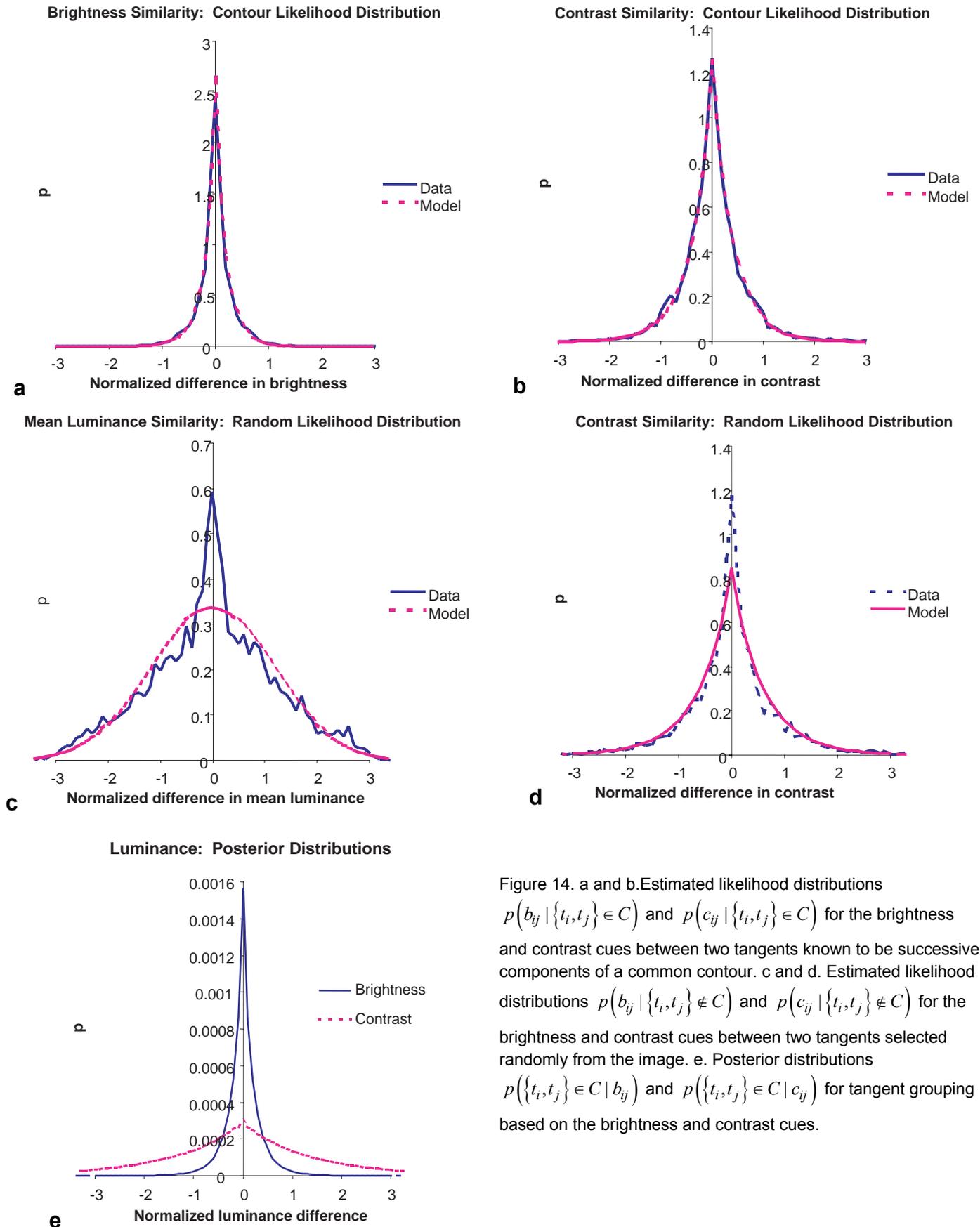


Figure 14. a and b. Estimated likelihood distributions $p(b_{ij} | \{t_i, t_j\} \in C)$ and $p(c_{ij} | \{t_i, t_j\} \in C)$ for the brightness and contrast cues between two tangents known to be successive components of a common contour. c and d. Estimated likelihood distributions $p(b_{ij} | \{t_i, t_j\} \notin C)$ and $p(c_{ij} | \{t_i, t_j\} \notin C)$ for the brightness and contrast cues between two tangents selected randomly from the image. e. Posterior distributions $p(\{t_i, t_j\} \in C | b_{ij})$ and $p(\{t_i, t_j\} \in C | c_{ij})$ for tangent grouping based on the brightness and contrast cues.

Figures 14a and 14b show the contour likelihood distributions $p(b_{ij} | \{t_i, t_j\} \in C)$ and $p(c_{ij} | \{t_i, t_j\} \in C)$ for the normalized brightness and contrast cues. Both are well modeled by generalized Laplacian distributions. Figures 14c and 14d show the random likelihood distributions $p(b_{ij} | \{t_i, t_j\} \notin C)$ and $p(c_{ij} | \{t_i, t_j\} \notin C)$ for these two luminance cues. Unfortunately, the generalized Laplacian models these data less accurately, failing to completely reflect the sharp peaks observed in the data near zero. This result may reflect long-range spatial correlations in intensity values, so that tangents on the same object typically generate much lower luminance cue values than tangents from different objects. The generalized Laplacian parameters for these models are provided in Table 4.

Table 4. Generalized Laplacian Parameters for Similarity Cue Likelihoods

	σ	Kurtosis	γ
Normalized brightness cue likelihood (contour)	0.32	4.40	0.87
Normalized contrast cue likelihood (contour)	0.59	3.3	0.96
Normalized brightness cue likelihood (random)	1.2	-0.03	2.0
Normalized contrast cue likelihood (random)	0.87	3.3	0.97

The derived posterior distribution models $p(\{t_i, t_j\} \in C | b_{ij})$ and $p(\{t_i, t_j\} \in C | c_{ij})$ for these two cues are shown in Figure 14c. It is clear that the brightness cue is a far more powerful cue than the contrast cue.

7. Discussion

7.1 Cue Independence

We have assumed that the cues under study are mutually independent when conditioned upon $\{t_i, t_j\} \in C$ or $\{t_i, t_j\} \notin C$. As a first step in checking this assumption, we computed the Pearson correlation coefficients between the absolute values of the cues in the contour condition (Table 5). We see from this calculation that correlations are relatively small (less than 0.1) except for the brightness/contrast correlation (0.53). However, the relatively weak inferential power of the contrast cue (see below) suggests that the contrast cue could be omitted from models of contour grouping without substantial loss in inferential power.

Table 5. Pearson Correlations Between Grouping Principles

Grouping cues	Parallelism	Cocircularity	Brightness	Contrast
Proximity	0.01	-0.10	0.09	0.07
Parallelism		-0.05	-0.04	0.08
Cocircularity			0.09	0.07
Brightness				0.53

Numbers represent correlations between the absolute values of the cues.

Although we have not in this study considered contours that reverse in contrast, it is interesting to note that there is recent psychophysical evidence for a lack of independence in contrast sign and good continuation cues. In an extension of prior work that examined the role of grouping cues in the visual search for fragmented two-dimensional shapes (Elder & Zucker, 1993), Spehar (2002) found that discontinuities in tangent orientation and contrast sign had greater impact when they occurred at the same point in the contour than when they occurred at different points.

7.2 Cue Generality

Our measurements estimate the statistics of the ensemble of natural images, but the human visual system must make correct inferences about single images and image sequences. It is therefore important to know the dispersion in these statistics over images. If these statistics vary wildly from image to image, general statistical models of these grouping principles are of limited value.

As a first step, we examined the dispersion in the standard deviation of our cues for the contour condition (Table 6). Standard deviations over images range from 8%-28%. These figures aid in interpreting the distributions estimated in this work. While knowledge of these distributions is clearly valuable, precision beyond roughly 8%-28% is of little value, because at this point errors due to image idiosyncrasies will begin to dominate. This observation also provides guidance in the search for correspondences between the statistics of the natural world and visual processing in humans. If we find agreements to within this accuracy, it is reasonable to claim that the visual system is matched to the statistics of the natural world.

Table 6. Dispersion of Standard Deviation Statistic for 5 Cues Over 9 Images

Grouping cue	Standard deviation over images
Proximity	28%
Parallelism	10%
Cocircularity	8%
Brightness	20%
Contrast	27%

7.3 Inferential Power

One of the main goals of this work is to estimate the inferential power of each of the contour grouping principles under consideration. Here we evaluate these principles using two different measures of inferential power. As we shall see, these two distinct measurements will produce a consistent ranking of the cues in terms of their statistical power.

7.3.1 Mutual Information

We want to measure the statistical power of each cue for the grouping of two tangents $\{t_i, t_j\}$. We define a random variable $G_{ij} \in \{G_1, G_2\}$ to represent the actual grouping relationship between tangents t_i and t_j , where G_1 represents the event $\{t_i, t_j\} \in C$ and G_2 represents the event $\{t_i, t_j\} \notin C$. Prior to making any observations about the relationship between the two tangents, there is an inherent uncertainty in the decision variable G_{ij} that can be formally characterized by its informational entropy $H(G_{ij})$ (Shannon & Weaver, 1949):

$$H(G_{ij}) = -p(G_1) \log_2 p(G_1) - p(G_2) \log_2 p(G_2). \quad (11)$$

If it were the case that two arbitrary tangents were as likely to be grouped as not, the prior entropy of the grouping decision would be exactly one bit. However, because it is far more likely that two arbitrary tangents are not grouped, the prior entropy is much lower. If we assume that, on average, a tangent groups with one other tangent in each direction,¹⁰ then $p(G_1) \approx 1/n$ and $p(G_2) \approx 1 - 1/n$, where n is the number of tangents in the image. For our sample of 9 images, in which the mean number of tangents per image is roughly 5,000, the prior entropy $H(G_{ij})$ of the grouping decision is roughly 0.0027 bits.

Grouping cues provide knowledge about the relationship between tangents that can reduce the uncertainty in the grouping decision. The size of this reduction in uncertainty can be characterized by the mutual information $I(G_{ij}; d_{ij})$ between the grouping decision G_{ij} and the cue d_{ij} :

$$I(G_{ij}; d_{ij}) = \int p(G_1, d_{ij}) \log_2 \frac{p(G_1, d_{ij})}{p(G_1)p(d_{ij})} dd_{ij} + \int p(G_2, d_{ij}) \log_2 \frac{p(G_2, d_{ij})}{p(G_2)p(d_{ij})} dd_{ij}. \quad (12)$$

Given observation of the grouping cue d_{ij} , the remaining uncertainty in the grouping decision G_{ij} is given by

$$H(G_{ij} | d_{ij}) = H(G_{ij}) - I(G_{ij}; d_{ij}). \quad (13)$$

Thus the mutual information measures the information, in bits, that the cue d_{ij} provides about the decision G_{ij} of whether to group tangents t_i and t_j . Normalizing the mutual information $I(G_{ij}; d_{ij})$ by the prior uncertainty $H(G_{ij})$ yields the percentage reduction in informational uncertainty provided by the cue. Using this measure, Figure 15 shows the informational power of each of the four cues under consideration. It is clear that the proximity cue is the most powerful, reducing the entropy in the grouping decision by 75%. The combined power of the good continuation cues appears to be

roughly comparable to the power of the similarity cues. We can also see that the parallelism cue is substantially more powerful than the cocircularity cue, and the brightness cue is much more powerful than the contrast cue.

Here we considered only grouping that preserved the polarity of contrast along the contour. If we were to generalize the study to allow contrast reversals, how much inferential power could be obtained by knowledge of the sign of contrast of two tangents?

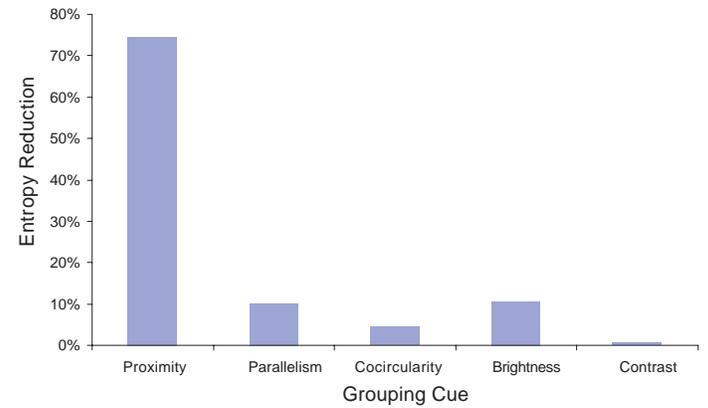


Figure 15. Inferential power of contour grouping cues, as measured by the proportion of entropy in the grouping decision eliminated by knowledge of the cue.

Our best guess is that reversals occur around 5% of the time. In a typical image of 5,000 tangents, this cue represents a reduction in entropy of about 5%. If reversals occur less frequently, the cue will be more powerful, and in the limit (when reversals never occur), the cue represents an entropy reduction of about 7%. The reason for this rather modest limit is that random tangents are as likely to be contrast reversing as not, so the cue tell us little about half of the possible local groupings in the image.

In summary, we expect the contrast sign cue to be less informative than the brightness cue, but more informative than the contrast cue. Together, these similarity cues may be slightly more powerful than the combined good continuation cues.

We find that the reduction in the decision entropy by proximity, good continuation, and similarity cues sums to roughly 100%. This implies that, were these cues all perfectly independent, they could be used to construct a perfect grouping system. Because they are not perfectly independent, some uncertainty will remain, even when the cues are optimally combined.

7.3.2 Local/Global Model

This study focuses on the statistics of local cues for contour grouping. We expect that the perceptual organization of natural images may depend also upon global cues or constraints (Section 3). One possible model is that local cues are used to generate multiple local

hypotheses, from which final organizations are selected using global criteria, e.g., endpoints (Elder & Goldberg, 2001), closure (Elder & Zucker, 1996a), or completeness (Elder & Krupnik, 2001). The cost of this computation depends on the number of local hypotheses generated per tangent; stronger local cues may permit a reduction in the number of local hypotheses needed to ensure that the correct global solution is represented.

This theory motivates a third measure of local inferential power, based on the concept of a local complexity limit. We characterize the local complexity limit of a computation as the maximum number m of local hypotheses that may be represented per tangent. A local representation is then said to be in error if it does not contain the correct hypothesis.

Figure 16 shows the error rates for each of the cues individually,¹¹ and for all cues combined, as a function of the number of local hypotheses represented. Again we see that proximity is a far more important cue than good continuation and similarity cues, which are roughly equal in their inferential power. Importantly, we also see that Bayesian combination of the cues (Equation 7) yields the best performance, attesting to the independent information available in each cue.

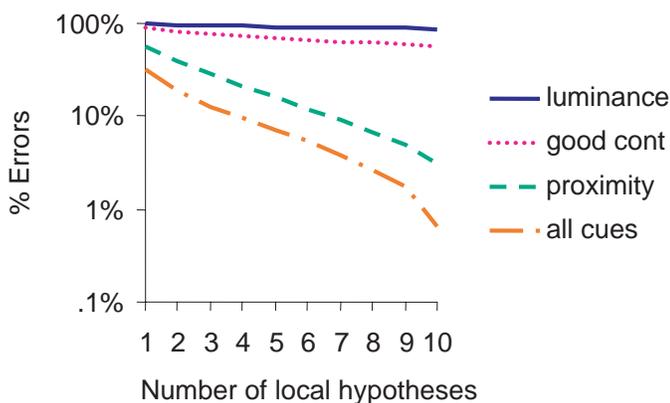


Figure 16. Contour grouping performance for each cue individually and for the combined cues.

7.4 On the General Shape of the Distributions

Contour likelihood distributions and posteriors for all cues were found to be kurtotic, with long tails. Thus extreme values for these cues occur as generic events. Figure 17 shows sample contours generated using the developed models for proximity, parallelism, and cocircularity cues. While these contours are generally continuous and smooth, sudden gaps and corners occur fairly frequently. This corresponds well to our intuition about contours in the natural world. Occlusions are commonplace, and thus we expect to occasionally observe significant gaps. Objects and other structures often have

corners, and thus we expect to occasionally witness significant jumps in tangent orientation. Objects can also be dappled in paint and shadow, and so discontinuities in luminance cues are also likely.

The generative nature of our model is quite important. Generated sample contours allow us to visually evaluate the statistical information captured in the model, and can be used in psychophysical experiments to assess the tuning of human vision to natural image statistics. Note that modeling the ordered statistics is critical for the generative model: the unordered statistics are insufficient.



Figure 17. Sample contours generated from natural image statistics.

7.5 Quantitative Models for the Law of Proximity

Given the relative importance of the law of proximity, it would be a significant advance if we could unify the statistics of proximity cues in natural images with observed psychophysical data. In Section 1.1, we noted that Gaussian, exponential, and power law models have been used to describe the law of proximity in both human and computer vision systems. Our results show that only the power law is consistent with the ecological statistics of contours. Power laws generally suggest scale-invariant processes, and thus our findings support prior psychophysical observations of scale invariance in the perceptual organization of dot lattice stimuli (Kubovy & Holcombe, 1998).

We can compare the estimated exponent $b = 2.92 \pm 0.02$ to estimates made by Oyama based on psychophysical data for multistable dot-lattice patterns (Oyama 1961; see Section 1.1). From two separate experiments, Oyama estimated $b = 2.88$ and $b = 2.89$. These psychophysical estimates differ by roughly 1% from our own estimate based on the statistics of natural images. This difference is not statistically significant ($p < .05$).

The similarity of our statistical measurements of natural image contours to these psychophysical results is intriguing, but we must keep in mind that this comparison assumes (1) that tangent grouping can be compared to dot grouping and (2) that the perception of a multistable percept is proportional to the probability of the configuration. These assumptions should be tested in future work.

Sigman et al. (2001) found that spatial correlation in the response of collinearly-oriented filters to natural images followed a power law with a much weaker exponent ($b = 0.6$).¹² The difference between exponents is not surprising, given the differences in the types of measurements made in these two studies. Recall that Sigman et al. measured correlations without regard to whether the image points were contiguous components of the same underlying contour, or indeed whether they were on the same contour or on a contour at all. We thus expect to observe in their statistics much longer-range correlations contributed by widely spaced points on contours and texture flows.

Our own approximate analysis of the data from Geisler et al. (2001) (Section 7.6) suggests a power law with an exponent of roughly 1.4, half the exponent derived from our data. Recall that Geisler and colleagues' statistics do pertain to tangents on the same contour, but do not reflect the ordering of the tangents on the contour. The exponent of 2.92 that we estimate from our own data reflects the power law relating directly successive tangents on a common contour. Thus we witness a progression of stronger proximity laws as we consider the statistics that are most relevant to the specific problem of contour grouping.

It is encouraging that our statistical results are consistent with the psychophysical results of Oyama, but it is unclear why other psychophysical studies suggest exponential or Gaussian distributions (Oeffelen & van Vos, 1982; Compton & Logan, 1993; Elder & Zucker, 1994; Kubovy & Wagemans, 1995; Kubovy & Holcombe, 1998). One possibility is that a power law simply was not tested for these experiments.¹³ It is also possible that the observed distribution may depend on the nature of the experiment. A full treatment of this issue is beyond the scope of this work, but it seems likely that a reexamination of past experiments in light of these new data on the ecological statistics of the proximity principle may lead to a more consistent quantitative understanding of this most powerful of Gestalt laws.

7.6 Association Fields

The role of proximity and good continuation cues in contour grouping can be vividly seen in Figure 18, where we have used the method from Geisler et al. (2001) for visualizing the association fields¹⁴ defining contour grouping. Figure 18a shows the relative probability and mean orientation of a tangent relative to the directly

preceding tangent on the contour, shown in black at the origin. Note the rapid falloff with separation, reflecting the power of the proximity cue, and the fact that tangents are more nearly parallel than cocircular, reflecting the greater power of the parallelism cue over the cocircularity cue. Similar observations can be made from Figure 18b and 18c, which show the likelihood ratio and posterior probability for each possible tangent position and relative orientation. Figure 18d shows the posterior probabilities determined from our factorized model combining proximity, parallelism, and cocircularity cues.

Figure 18e shows the likelihood ratios estimated by Geisler et al. (2001). These should be compared to our estimates in Figure 18b. Several differences can be observed. First note that because we encode the topology of the contours, our association fields are one-sided, i.e., they reflect grouping to the rightward end of the tangent shown in black at the origin. Second, there is a difference in scale. We plot our association fields on a log scale, because of the underlying power law governing the proximity cue. While the minimum and maximum separation plotted for our data are 1.4 and 23 pixels, the minimum and maximum plotted by Geisler et al. are roughly 3.7 and 41 pixels.

Next, note that the likelihood ratios we measure are as high as 10^6 , whereas those measured by Geisler et al. (2001) are less than 10^2 . We believe that this difference is due primarily to two factors. First, because of the power law governing the proximity cue, the likelihood ratio is very high for very small separations, and because we report probabilities for smaller separations, higher likelihood ratios are to be expected. Second, in order to interpret this difference, we must convert the likelihood ratios to posterior probabilities, which also depend on the prior ratios (Equation 1, Section 4). In our data, the prior ratio is equal to the number of tangents in the average image, whereas in the data from Geisler et al., the prior ratio is approximately equal to the average number of distinct contours traced in an image.

More interesting is the difference in the shape of the probability surface. The probability surface based on our data falls off much more quickly with separation than the probability surface based on the Geisler et al. (2001) data. For example, in our statistics, the likelihood ratio falls by a factor of roughly 100 when separation is increased by a factor of 4.7 from 3.6 pixels to 17 pixels, consistent with the power law exponent of 2.88 we estimated in Section 6.1. In contrast, in the Geisler et al. statistics, the likelihood ratio falls by a factor of only 10 when separation is increased by a factor of 5, from 3.7 to 18.6 pixels, suggesting a power law exponent of about 1.44, half the value we obtain. Thus the proximity cue is far more powerful for grouping contours in correct topological order than for grouping contours as unordered sets of tangents. This explains, at least in part, why the principal of proximity was not emphasized in other studies.

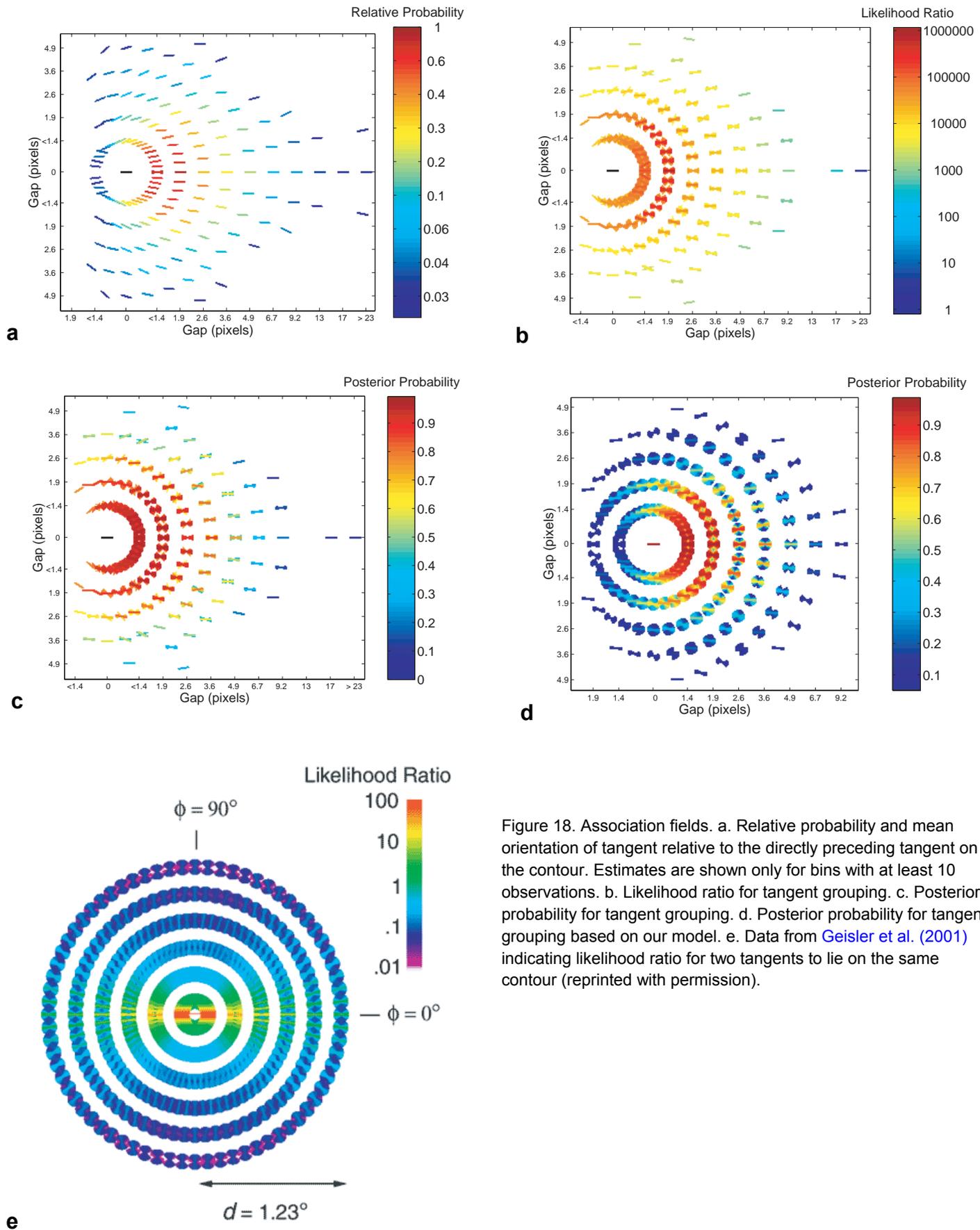


Figure 18. Association fields. a. Relative probability and mean orientation of tangent relative to the directly preceding tangent on the contour. Estimates are shown only for bins with at least 10 observations. b. Likelihood ratio for tangent grouping. c. Posterior probability for tangent grouping. d. Posterior probability for tangent grouping based on our model. e. Data from Geisler et al. (2001) indicating likelihood ratio for two tangents to lie on the same contour (reprinted with permission).

It is less clear why our data overweight the parallelism cue and underweight the cocircularity cue relative to the Geisler et al. data. One possibility is the difference in the image databases used. Whereas Geisler et al. (2001) used strictly natural images, all but two of the images we used contained man-made structures. Because the human visual system develops in a mixture of these two environments, it would be useful to conduct a more controlled comparative study of their statistics.

8. Summary of Results and Relationship to Other Studies

Here we summarize the results of this work in the context of related studies.

8.1 Defining the Problem

Kruger (1998), Geisler et al. (2001), and Sigman et al. (2001) examined the joint statistics of oriented filter responses in natural images, and found statistical correlations suggestive of proximity, colinearity, parallelism, and cocircularity principles. However, no attempt was made to restrict the analysis to filters responding to contours, let alone common contours. These statistics therefore arise from within and between a mixture of structures: contours, texture, and shading flows, and do not lead directly to a model for the perceptual organization of contours.

Geisler et al. (2001) have taken this work one step further. Using human participants to trace perceived contours, the second-order spatial statistics of oriented elements lying on a common contour are estimated. However, the contour tracing method employed by Geisler et al. does not recover the ordering of the elements along a contour. Thus a contour is represented as an unordered set of oriented elements, and the statistics relate arbitrary pairs of tangents on a contour.

We have argued that a statistical encoding of the ordering of elements along a contour is critical to understanding contour grouping. A defining property of contours is their one-dimensionality. Without this, basic properties of contours, such as curvature, closure, concavity, and convexity, cannot be defined. In continuous space, contours may be parameterized by a real variable, e.g., arc length. To maintain this property in a discrete encoding, a contour must be represented not as a set but as a sequence of local elements.

For these reasons, we define the problem of contour grouping as the recovery of sequences of tangents projecting from the contours of a scene. We model these sequences as Markov chains. The Markov approximation captures the local nature of the physical processes that give rise to these contours, and is consistent with the monotonically decreasing nature of the autocorrelation function of natural images, with the spatiotopic structure

of early visual cortex and with the well studied psychophysical principle of proximity. Application of the Markov approximation allows us to understand the problem of contour grouping by characterizing the statistics of local grouping between successive elements comprising a contour.

Our model reflects the fact that the statistical dependencies between neighbouring tangents on a contour are much stronger than those between distant tangents on a contour. In the approach of Geisler et al. (2001), the power of the strong statistics relating neighboring tangents is diluted with the weak statistics relating distant tangents. This leads to substantial differences between our statistics and their statistics.

8.2 Identification of Roughly Independent Grouping Cues

Gestalt psychologists identified a number of distinct grouping principles, including proximity, good continuation, and similarity. They understood that these cues acted together to determine a percept, but the fact that they were (and still are) distinguished as separate cues in the literature suggests that their interaction may be relatively simple. In modern language, we might hope that these cues are relatively independent, and thus can be modeled by their marginal distributions, which can be combined factorially.

Whereas other studies (Kruger, 1998; Sigman et al., 2001; Geisler et al., 2001) have primarily reported results as multidimensional histograms, our study has focused on the recoding of observable variables into cues that are approximately independent. This process is important, for it allows us to better understand the structure of the multidimensional distributions describing natural contours.

Here we have shown that, to a first approximation, proximity, parallelism, cocircularity, brightness similarity, and contrast similarity can be considered as independent cues to contour grouping.

8.3 Parametric Modeling

Whereas other studies have primarily reported results graphically, in our study we have focused on developing accurate parametric models for grouping cues.

The development of parametric models is important for a number of reasons. First, they summarize a vast quantity of data in a small representation that is freely available for comparison with other studies, for use in psychophysical experiments, computational models, and computer vision algorithms. For example, the representation of the proximity cue by a power law allowed us to easily compare our results to the psychophysical results of Oyama (1961). Second, interpolation or extrapolation from these models leads to testable statistical predictions for particular contour

configurations that may not have been observed in our study but may be of interest in later studies. Third, the functional form of a model that accounts well for the data provides insight into the underlying physical process. For example, the power law discovered to account for the proximity cue is suggestive of a scale-invariant process. The general finding that our distributions are kurtotic predicts contours with sudden gaps and corners (Figure 17).

8.4 Proximity Cue

Our finding that the proximity cue follows a power law suggestive of a scale-invariant physical process is consistent with power laws found for simpler spatial statistics of natural images (Field, 1987; Ruderman & Bialek, 1994). The estimated exponent of this power law (2.92) is very close to that determined psychophysically in previous dot lattice experiments (Oyama, 1961).

Interestingly, Sigman et al. (2001) also report a power law for the proximity of colinear elements, with a much weaker exponent of 0.6.¹² Our own approximate analysis of the data of Geisler et al. (2001) (Section 7.6) suggests a power law with an exponent of 1.4, roughly half the exponent for our data. Thus we witness a progression of stronger proximity laws as we converge to the statistics that are most relevant to contour grouping.

8.5 Good Continuation Cue

We introduced a linear recoding of good continuation cues into parallelism and cocircularity cues. This recoding transforms the cues into an intuitive, roughly independent basis set. Another advantage of partitioning the good continuation cue into parallelism and cocircularity cues is that the sources of error determining the accuracy with which these cues can be estimated are quite different. Whereas the parallelism cue depends (linearly) on only estimation of tangent orientation, the cocircularity cue also depends (nonlinearly) on estimation of tangent endpoint position. This dependency makes the estimation of the cocircularity cue ill posed for closely spaced tangents. We showed that if this source of error is accounted for, both the parallelism and cocircularity cues are approximately uncorrelated with the proximity cue.

8.6 Similarity (Intensity) Cues

Whereas other studies have focused solely on geometric cues to contour grouping, we have also examined photometric cues. We have shown that these cues are comparable in inferential power to the good continuation cues (see below).

8.7 Inferential Power

Partitioning of the observable information relevant to contour grouping into distinct, roughly independent cues permits the evaluation of the relative inferential power of each of these cues, something not done in other studies. Our finding that proximity is by far the most powerful cue for contour grouping runs counter to prior empirical studies that have emphasized parallelism and colinearity (Kruger, 1998), and both theoretical and empirical work that has emphasized cocircularity (Parent & Zucker, 1989; Zucker, Dobbins, & Iverson, 1989; Sigman et al., 2001; Geisler et al., 2001).

Other empirical studies have not explicitly measured the relative power of these distinct cues, and this may partly explain the difference in emphasis. However, it is likely that the role of proximity was indeed weaker in these studies because the topology of the contour was being ignored. In other words, tangents that are adjacent in the sequence of elements defining a contour are likely to be closer together than tangents sampled randomly from the contour, and this will result in a more powerful role for the proximity cue.

We find photometric cues to be comparable in their importance to good continuation cues, which is interesting because their role in contour grouping has been much less studied. Brightness was found to be a far more powerful cue than contrast. A theoretical analysis of contrast polarity suggests that this could be a significant cue, less important than parallelism or brightness, but more important than cocircularity or contrast.

8.8 A Generative Model for Contour Grouping

Our model for contour grouping is generative: sample contours can be constructed from our models for the likelihood distributions of each cue (Figure 17). This is useful for visualizing various implications of the model, and for generating naturalistic stimuli for psychophysical studies. Note that there is no obvious way to use the statistics reported by other studies (Kruger, 1998; Sigman et al., 2001; Geisler et al., 2001) to generate sample contours, because the ordinal structure of the contours are not represented in these statistics.

8.9 A Bayesian Algorithm for Contour Grouping

Our study leads directly to a well-defined parameter-free Bayesian algorithm for contour grouping. The Markov approximation means that the joint probability of a contour can be expressed as a product of local grouping probabilities. Given a global goal (e.g., find the most probable contour of a specified length, find the most probable closed contour passing through a specified tangent, find the most probable contour between two

specified tangents), the optimal solution is well defined, with no free parameters. All of these problems can be solved using a shortest path algorithm in polynomial time (Elder & Zucker, 1996a; Elder & Goldberg, 2001). Application of other global constraints (e.g., simplicity [no self-intersections]) may lead to a more complex problem best solved approximately using probabilistic search techniques (Elder & Krupnik, 2001).

Geisler et al. (2001) propose a simple algorithm based on a transitivity heuristic and an adjustable grouping threshold parameter. Neither the transitivity heuristic nor the setting of the grouping threshold has justification in terms of statistical theory. The algorithm has not, to our knowledge, been demonstrated on natural images.

9. Conclusions

We have studied the statistics of contour grouping cues in natural images. We considered three of the classical Gestalt principles of perceptual organization: proximity, good continuation and similarity (in brightness and contrast). We introduced a technique for rapidly collecting these statistics based upon contour tracing by human observers. We found that to a first approximation these cues can be considered independent. The dispersion in these statistics over our sample of images indicates that approximate knowledge of these statistics is useful for statistical inference in single images. We found large differences in the statistical power of these cues for contour grouping; the law of proximity is by far the most powerful. Quantitative modeling indicates that the proximity cue for contour grouping follows a power-law with exponent in close agreement with prior psychophysical data on the perception of multistable dot lattice displays. It is hoped that these statistics will lead to the discovery of further links between the perceptual organization of contours and the statistics of the natural world.

10. Acknowledgments

We thank Bill Geisler and Bruno Olshausen for helpful comments on the manuscript. We thank Pat Bennett for the reference to Brunswik and Kamiya (1953). This research was supported in part by grants from the Natural Sciences and Engineering Research Council of Canada and the Geomatics for Informed Decisions Network of Centres of Excellence. Commercial relationships: None.

11. Footnotes

¹ In fact, efficiency was found to decline slightly with elongation, but this could be due to a decrease in acuity with eccentricity.

² An exponential decay model has also been used in stochastic computer vision models of contours (Mumford, 1992; Williams & Jacobs, 1997).

³ Although mean density was roughly equated, elements were more evenly spaced along the contour, so that a regularity cue cannot be completely ruled out.

⁴ See Simoncelli & Olshausen (2001) for an excellent review.

⁵ We do not consider edge blur as a grouping cue in this study.

⁶ We test this hypothesis in Section 7.1.

⁷ It is hoped that the present study will provide the data for a more principled choice of distributions and parameters.

⁸ Since a typical image contains on the order of 5,000 tangents, each image contains on the order of 25 million tangent pairs.

⁹ In fact, the parallelism and cocircularity cues are restricted to within $\pm 360^\circ$, so we use a truncated form of the generalized Laplacian distribution for these good continuation cues.

¹⁰ This is a statement of the expected case, and does not preclude contour terminations, where a tangent does not group at all in one direction, nor bifurcations, where a tangent groups with more than one other tangent in one direction.

¹¹ We have combined the two good continuation cues and the two similarity cues in this graph.

¹² The fact that a power law $p(x) = ax^{-b}$ is not integrable over $[x_0, \infty)$ for $b \leq 1$ suggests that the data of Sigman et al. (2001) are not well-modeled by a power law.

¹³ Kubovy and Wagemans (1995) do note that a power law would fit their data equally well, but suggest that there is no theoretical reason to favour one distribution over the other.

¹⁴ The term “association field” was coined by Field et al. (1993), to refer to a diagram indicating the rules by which elements in a path “are associated and segregated from the background”. Similar diagrams can be found in Parent and Zucker (1989) and Zucker, Dobbins, and Iverson (1989).

12. References

- Alter, T. D. (1995). The role of saliency and error propagation in visual object recognition. (Doctoral dissertation, MIT, 1995). *Dissertation Abstracts International*, 56, 3844.
- Atick, J. J., & Redlich, A. N. (1992). What does the retina know about natural scenes? *Neural Computation*, 4, 196-210.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, 61, 183-193.
- Barlow, H. B. (1961). The coding of sensory messages. In W. Thorpe & O. Zangwill (Eds.), *Current problems in animal behavior* (pp. 331-360). Cambridge, UK: Cambridge University Press.

- Barlow, H. B. (1978). The efficiency of detecting changes of density in random dot patterns. *Vision Research*, 18, 637-650. [PubMed]
- Beck, J., Rosenfeld, A., & Ivry, R. (1989). Line segregation. *Spatial Vision*, 4, 75-101. [PubMed]
- Brunswik, E., & Kamiya, J. (1953). Ecological cue-validity of 'proximity' and of other Gestalt factors. *American Journal of Psychology*, 66, 20-32.
- Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, 197, 551-566. [PubMed]
- Compton, B. J., & Logan, G. D. (1993). Evaluating a computational model of perceptual grouping by proximity. *Perception and Psychophysics*, 53, 403-421. [PubMed]
- Cox, I. J., Rehg, J. M., & Hingorani, S. (1993). A Bayesian multiple-hypothesis approach to edge grouping and contour segmentation. *International Journal of Computer Vision*, 11, 5-24.
- Crevier, D. (1999). A probabilistic method for extracting chains of collinear segments. *Computer Vision and Image Understanding*, 76, 36-53.
- David, C., & Zucker, S. W. (1990). Potentials, valleys, and dynamic global coverings. *International Journal of Computer Vision*, 5, 219-238.
- DeValois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, 22, 545-559. [PubMed]
- Dudek, G., & Tsotsos, J. K. (1997). Shape representation and recognition from curvature. *Computer Vision, Graphics and Image Processing: Image Understanding*, 68, 170-189.
- Earle, D. C. (1999). Glass patterns: Grouping by contrast similarity. *Perception*, 28, 1373-1382. [PubMed]
- Elder, J. H. (1999). Are edges incomplete? *International Journal of Computer Vision*, 34, 97-122.
- Elder, J. H., & Goldberg, R. M. (1998a). Inferential reliability of contour grouping cues in natural images. *Perception*, 27, 11.
- Elder, J. H., & Goldberg, R. M. (1998b). *The statistics of natural image contours*. Paper presented at the 1998 IEEE Computer Society Workshop on Perceptual Organization in Computer Vision. Available at marathon.csee.usf.edu/~sarkar/pocv_program.html.
- Elder, J. H., & Goldberg, R. M. (2001). Image editing in the contour domain. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23, 291-296.
- Elder, J. H., & Krupnik, A. (2001). Contour grouping with strong prior models. In *Proceedings of the IEEE Conference on Computer Vision Pattern Recognition* (pp. 414-421). Kauai, Hawaii: IEEE Computer Society Press.
- Elder, J., & Zucker, S. (1993). The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision Research*, 33, 981-991. [PubMed]
- Elder, J., & Zucker, S. (1994). A measure of closure. *Vision Research*, 34, 3361-3370. [PubMed]
- Elder, J. H., & Zucker, S. W. (1996a). Computing contour closure. In *Proceedings of the 4th European Conference on Computer Vision* (pp. 399-412). New York: Springer Verlag.
- Elder, J. H., & Zucker, S. W. (1996b). Scale space localization, blur and contour-based image coding. In *Proceedings of the IEEE Conference on Computer Vision Pattern Recognition* (pp. 27-34). San Francisco, CA: IEEE Computer Society Press.
- Elder, J. H., & Zucker, S. W. (1998a). Evidence for boundary-specific grouping. *Vision Research*, 38, 143-152. [PubMed]
- Elder, J. H., & Zucker, S. W. (1998b). Local scale control for edge detection and blur estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 699-716.
- 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. [PubMed]
- 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, 173-193. [PubMed]
- Freeman, W. T. (1992). Steerable filters and local analysis of image structure (Doctoral dissertation, MIT, 1992). *Dissertation Abstracts International*, 53, 4846.
- Freeman, W. T., & Adelson, E. H. (1991). The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13, 891-906.
- Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, 41, 711-724. [PubMed]

- Geisler, W. S., Super, B., & Gallogly, D. (2000). Natural image statistics predict contour detection performance [Abstract]. *Investigative Ophthalmology and Visual Science*, *41*, 1672.
- Gilbert, C. D., & Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *Journal of Neuroscience*, *9*, 2432-2443. [PubMed]
- Glass, L., & Switkes, E. (1976). Pattern recognition in humans: Correlations which cannot be perceived. *Perception*, *5*, 67-72. [PubMed]
- Hochberg, J. (1974). Organization and the Gestalt tradition. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (pp. 179-210). New York: Academic Press.
- Hochberg, J., & Hardy, D. (1960). Brightness and proximity factors in grouping. *Perceptual and Motor Skills*, *10*.
- Huttenlocher, D. P., & Wayner, P. C. (1992). Finding convex edge groupings in an image. *International Journal of Computer Vision*, *8*, 7-27.
- Jacobs, D. W. (1996). Robust and efficient detection of salient convex groups. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *18*, 23-37.
- Kass, M., Witkin, A., & Terzopoulos, D. (1987). Snakes: Active contour models. *International Journal of Computer Vision*, *1*, 321-331.
- Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt, Brace, & World.
- Kruger, N. (1998). Collinearity and parallelism are statistically significant second order relations of complex cell responses. *Neural Processing Letters*, *8*, 117-129.
- Kubovy, M., & Holcombe, A. O. (1998). On the lawfulness of grouping by proximity. *Cognitive Psychology*, *35*, 71-98. [PubMed]
- Kubovy, M., & Wagemans, J. (1995). Grouping by proximity and multistability in dot lattices: A quantitative Gestalt theory. *Psychological Science*, *6*, 225-234.
- Lowe, D. G. (1989). Organization of smooth image curves at multiple scales. *International Journal of Computer Vision*, *3*, 119-130.
- Mahamud, S., Thornber, K. K., & Williams, L. R. (1999). Segmentation of salient closed contours from real images. In *Proceedings of the 7th IEEE International Conference on Computer Vision* (pp. 891-897). Los Alamitos, CA: IEEE Computer Society Press.
- Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, *11*, 674-693.
- Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings of the 8th IEEE International Conference on Computer Vision* (pp. 416-424). Los Alamitos, CA: IEEE Computer Society Press.
- Mumford, D. (1992). Elastica and computer vision. In C. Bajaj (Ed.), *Algebraic geometry and applications* (pp.491-506). Heidelberg, Germany: Springer-Verlag.
- Oeffelen, M. P., & van Vos, P. G. (1982). Configurational effects on the enumeration of dots: Counting by groups. *Memory and Cognition*, *10*, 396-404. [PubMed]
- Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, *381*, 607-609. [PubMed]
- Olshausen, B. A., & Field, D. J. (1997). Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision Research*, *37*, 3311-3325. [PubMed]
- Oyama, T. (1961). Perceptual grouping as a function of proximity. *Perceptual and Motor Skills*, *13*, 305-306.
- Parent, P., & Zucker, S. W. (1989). Trace inference, curvature consistency, and curve detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*, 823-839.
- Parker, A. J., & Hawken, M. J. (1988). Two-dimensional spatial structure of receptive fields in monkey striate cortex. *Journal of the Optical Society of America A*, *5*, 598-605. [PubMed]
- Perona, P. (1995). Deformable kernels for early vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *17*, 488-499.
- Ruderman, D. L., & Bialek, W. (1994). Statistics of natural images: Scaling in the woods. *Physical Review Letters*, *73*, 814-817. [PubMed]
- Sachs, A., & Elder, J. H. (2000). Estimating the psychophysical receptive fields of edge detection mechanisms. *Perception*, *29*, 122.
- Saund, E. (1990). Symbolic construction of a 2-D scale-space image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*, 817-830.

- Sha'ashua, A., & Ullman, S. (1988). Structural saliency: The detection of globally salient structures using a locally connected network. In *Proceedings of the 2nd International Conference on Computer Vision* (pp. 321-327). Washington: IEEE Press.
- Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.
- Sigman, M., Cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: Natural scenes and Gestalt rules. *Proceedings of the National Academy of Sciences*, *98*, 1935-1940. [PubMed]
- Simoncelli, E. P. (1997). Statistical models of images: Compression, restoration and synthesis. In *31st Asilomar Conference on Signals, Systems and Computers*. Los Alamitos, CA: IEEE Computer Society Press.
- Simoncelli, E. P. (1999). Modeling the joint statistics of images in the wavelet domain. *Proceedings of the SPIE 44th Annual Meeting* 3813, 188-195.
- Simoncelli, E. P., & Adelson, E. H. (1996). Noise removal via Bayesian wavelet coring. In *3rd IEEE International Conference on Image Processing* (pp. 379-382). Piscataway, NJ: IEEE Signal Processing Society.
- Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, *24*, 1193-1216. [PubMed]
- Spehar, B. (2002). The role of contrast polarity in perceptual closure. *Vision Research*, *42*, 343-350. [PubMed]
- Srinivisan, M. V., Laughlin, S. B., & Dubs, A. (1982). Predictive coding: A fresh view of inhibition in the retina. *Proceedings of the Royal Society of London B: Biological Sciences*, *216*, 427-459. [PubMed]
- Uttal, W. R., Bunnell, L. M., & Corwin, S. (1970). On the detectability of straight lines in visual noise: An extension of French's paradigm into the millisecond domain. *Perception and Psychophysics*, *8*, 385-388.
- Watson, A. B. (1982). Summation of grating patches indicates many types of detector at one retinal location. *Vision Research*, *22*, 17-25. [PubMed]
- Watson, A. B. (2000). Visual detection of spatial contrast patterns: Evaluation of five simple models. *Optics Express*, *6*, 12-33. [Link]
- Watt, R. J., & Morgan, M. J. (1984). Spatial filters and the localization of luminance changes in human vision. *Vision Research*, *24*, 1387-1397. [PubMed]
- Wertheimer, M. (1938). Laws of organization in perceptual forms. In W. D. Ellis (Ed.), *A sourcebook of Gestalt psychology* (pp. 71-88). London, UK: Routledge and Kegan Paul. (*Psychol. Forschung*, 1923, 4, 301-350)
- Williams, L. R., & Jacobs, D. W. (1997). Stochastic completion fields: A neural model of illusory contour shape and salience. *Neural Computation*, *9*, 837-858. [PubMed]
- Wilson, H. R., & Bergen, J. R. (1979). A four mechanism model for threshold spatial vision. *Vision Research*, *19*, 19-32. [PubMed]
- Wilson, H. R., & Gelb, D. J. (1984). Modified line-element theory for spatial-frequency and width discrimination. *Journal of the Optical Society of America A*, *1*, 124-131. [PubMed]
- Zucker, S. W., & Davis, S. (1988). Points and endpoints: A size/spacing constraint for dot grouping. *Perception*, *17*, 229-247. [PubMed]
- Zucker, S. W., Dobbins, A., & Iverson, L. (1989). Two stages of curve detection suggest two styles of visual computation. *Neural Computation*, *1*, 68-81.
- Zucker, S. W., Hummel, R., & Rosenfeld, A. (1977). An application of relaxation labeling to line and curve enhancement. *IEEE Transactions on Computing*, *26*, 394-403.
- Zucker, S. W., Stevens, K. A., & Sander, P. (1983). The relation between proximity and brightness similarity in dot patterns. *Perception and Psychophysics*, *34*, 513-522. [PubMed]