Communicated by Steven Zucker

Stochastic Completion Fields: A Neural Model of Illusory Contour Shape and Salience

Lance R. Williams
David W. Jacobs
NEC Research Institute, Princeton, NJ 08540, USA

We describe an algorithm- and representation-level theory of illusory contour shape and salience. Unlike previous theories, our model is derived from a single assumption: that the prior probability distribution of boundary completion shape can be modeled by a random walk in a lattice whose points are positions and orientations in the image plane (i.e., the space that one can reasonably assume is represented by neurons of the mammalian visual cortex). Our model does not employ numerical relaxation or other explicit minimization, but instead relies on the fact that the probability that a particle following a random walk will pass through a given position and orientation on a path joining two boundary fragments can be computed directly as the product of two vector-field convolutions. We show that for the random walk we define, the maximum likelihood paths are curves of least energy, that is, on average, random walks follow paths commonly assumed to model the shape of illusory contours. A computer model is demonstrated on numerous illusory contour stimuli from the literature.

1 Introduction

The completion shape and salience problem is the problem of computing the shape and relative likelihood (as determined by the prior distribution) of the family of curves that potentially connect (or complete) a set of boundary fragments. This is a necessary intermediate step in the solution of the full figural completion problem, which has been previously characterized (Williams & Hanson, 1996) as the problem of building a Huffman-labeled figure representing the boundaries of hidden and visible surfaces (see Figure 1). The fundamental assumption underlying our work is that the paths followed by a particle undergoing a random walk in a lattice of discrete positions and orientations can be used to model the prior probability distribution of the shape of boundary completions (see also Mumford, 1994). We observe that the activity of a population of neurons with receptive fields tuned to discrete positions and orientations (like early areas of primate visual cortex) can be interpreted as a probability distribution describing a particle's possible location (i.e., the current state of the Markov process defining the

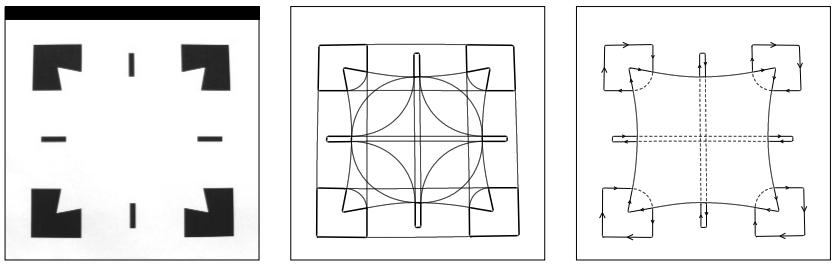

Figure 1: Williams and Hanson (1996) have characterized figural completion as the problem of building a Huffman-labeled figure from a set of boundary fragments. A Huffman-labeled figure consists of a set of closed, oriented plane curves (representing surface boundary components) together with an explicit indication of relative depth at crossing points. The labeling scheme is based on Huffman's concept of "X-ray pictures" of smooth objects (Huffman, 1971). (Left) Input stimulus. (Middle) Boundary fragments (thick) and potential boundary completions (thin) represented by cubic Bezier splines of least energy. The stochastic completion field (introduced in this article) is the corresponding parallel distributed representation. (Right) The human percept, represented as a Huffman-labeled figure. The problem of deriving the corresponding parallel distributed representation is a subject for future research.

random walk). Let there exist two subsets of states, P and Q, representing the beginning and ending points of a set of boundary fragments. We call P the set of *sources* and Q the set of *sinks.* Our goal is to compute the probability that a particle, initially in state (x_p, y_p, θ_p), for some $p \in P$, will in the course of a random walk (representing a prior on completion shape) pass through state (u, v, ϕ), on its way to state (x_q, y_q, θ_q), for some $q \in Q$, for all combinations of u, v, and ϕ. We call this parallel distributed representation of the family of curves that potentially connect pairs of boundary fragments a *stochastic completion field.*

To appreciate better the scope of this article, it will be useful to identify those phenomena for which we believe the stochastic completion field model can and cannot account. Our primary claim is that the mode, magnitude, and variance of the stochastic completion field are related to the shape, salience, and sharpness of perceived illusory contours. However, the definition of stochastic completion field says nothing about the specific forms of brightness stimuli that can elicit illusory contours. In particular, we do not describe a comprehensive theory of how raw image brightnesses are transformed into sources and sinks for our diffusion process. Furthermore, the stochastic completion field says nothing about apparent brightness, which is distinct from salience. Finally, we note that whether a completion is ac-

tually perceived cannot be solely a function of local configurational factors, but also depends on the role the completion plays in the larger surface organization (Williams & Hanson, 1996). Specifically, it depends on (1) whether it can be incorporated in a consistent way into a Huffman-labeled figure and (2) whether it is nominally visible (*modal*) or occluded (*amodal*). Neither of these can be determined by a process that does not consider the topology of the scene. Consequently, the stochastic completion field can only be regarded as a precursor to a true surface representation.

2 Previous Work

The novelty of our method lies in the definition of the stochastic completion field, which explicitly embodies the assumption that the prior distribution of completion shape can be modeled as a random walk. However, other researchers have advanced algorithm- and representation-level theories of figural completion based on orientation fields of various kinds. At the algorithm and representation level, all of these models are outwardly similar. This is because all derive from similar views of the orientation preference structure of the visual cortex, common assumptions about neural computation, and basic considerations (whether explicit or implicit) like translation and rotational invariance.

The earliest theory of illusory contour shape is due to Ullman (1976). Ullman hypothesized that the curve used by the human visual system to join two contour fragments is constructed from two circular arcs. Each circular arc is tangent to its sponsoring contour at one end and to the other arc at the point of intersection. From the family of possible curves of this form, the pair that minimizes total bending energy ($E = \int \kappa(s)^2 ds$ where $\kappa(s)$ is curvature parameterized by arc length) models the shape of the illusory contour. Ullman suggested that illusory contours could be computed in parallel in a network by means of local operations, but he did not implement or test this idea.

Grossberg and Mingolla (1985) describe a model of illusory contour formation that involves repeated convolution with a large-kernel filter. In outward form, this kernel resembles ours, but it does not represent the Green's function of a stochastic process of the sort described by us or Mumford (1994). The outputs of oppositely oriented filters are combined through a nonlinear operation consisting of thresholding followed by a logical AND function. Grossberg and Mingolla's network is complex, and its convergence properties are difficult to analyze. Clearly, part of this complexity stems from their desire to model, in a comprehensive way, the many different forms of stimulus that can elicit illusory contours (e.g., contrast edges of mixed sign, line endings). No other model (including ours) attempts to be this comprehensive.

Parent and Zucker (1989) describe a network algorithm for computing a consistent discrete tangent and curvature field from noisy local tangent and

curvature measurements. Their network solves a well-defined relaxation labeling problem where the goal is to find the most probable assignment of labels subject to a compatibility relation. The labels are discrete tangent and curvature values at every image point, and the compatibility relation is based on co-circularity. Although they do not claim to model illusory contour shape, David and Zucker (1990) use active minimum-energy-seeking contours to locate the valleys (or, equivalently, the mode or ridge lines) of a potential function derived from the output of the Parent and Zucker network. The ridge lines of the potential function represent the integral curves of the tangent field.

Like us, Heitger and von der Heydt (1993) describe a theory of figural completion based on nonlinear combination of the convolutions of "keypoint" images with a fixed set of oriented grouping filters. Significantly, they demonstrate their method on both illusory contour figures like the Kanizsa triangle and on more "realistic" images (e.g., of plants and rocks), with impressive results. Unfortunately, neither the equations defining the filters nor the proposed method of combination are described as a means of computing an underlying function. This makes analysis of their method difficult. However, because their grouping filters are scalar functions, they cannot (even implicitly) model a prior probability distribution of completion shapes in the manner we describe.

Guy and Medioni (1996) have described a method for computing a vector-field representing global image structure from local tangent measurements. Like the Hough transform, the key to their approach is the local summation of a set of global voting patterns. Unlike the Hough transform, the accumulator is spatially registered with the image, and the voting pattern is a vector-field, not a scalar-field. However, because the voting pattern represents orientations that are co-circular to the tangent measurements, the vector-field is nonstochastic (i.e., deterministic), and cannot model the prior distribution of completion shapes.

Shashua and Ullman (1988) describe a network algorithm for computing "perceptual saliency," although they do not portray it as a faithful model of human visual processing. Computing salience (as they define it) requires finding the energy of the minimum energy curve of length n beginning at every position and orientation. This energy function combines a term similar to squared curvature and a term measuring gap size. However, because the energy function also includes terms designed to guarantee convergence, the output of their system can be difficult to anticipate. Alter and Basri (1996) analyze the behavior of this network in detail and point out some of its counterintuitive behavior.

3 Fundamental Assumption

Like many other problems in vision, figural completion is ill posed; the visual system cannot know the precise shape of an object's boundary where

it is hidden from view by another object. The best that it can do in such a situation is to make an educated guess. It is conventionally assumed that this guess takes the form of a maximum likelihood estimate. That is, given a prior probability distribution of completion shapes, it is assumed that the visual system chooses the member of the distribution that is most probable. In this article, we propose a somewhat different approach. Instead of computing (just) the maximum likelihood completion shape, we propose that the visual system computes local image plane statistics of the distribution of possible completion shapes. There is more information contained in a distribution than just the location of the mode, and we believe that in the case of the distribution of boundary completion shapes, these other properties are psychophysically meaningful. In particular, we believe that the variance of the distribution about the mode is related to illusory contour sharpness.

Like Mumford (1994), we characterize the distribution of completion shapes using the mathematical device of a particle undergoing a stochastic motion (a directional random walk). The random walk embodies the Gestalt principles of proximity and good continuation, which we believe originate in the statistics of the environment our visual systems evolved in. However, apart from shortness and smoothness (which have their origins in the world), we make one additional assumption, the origin of which (we believe) is in our heads. This is the assumption that the distribution of shapes that can extend a contour at any point is totally determined by the position and orientation of the contour at that point, that is, that the distribution of shapes can be modeled as a Markov process.

An immediate consequence of this assumption is that for the random walk we define, it can be shown (see the appendix) that the maximum likelihood path followed by a particle joining two points at fixed orientations is a curve of least energy (where energy is a linear combination of the integral of curvature squared and length). This is the curve that, others have theorized (Horn, 1981), models the shape of illusory contours joining boundary fragments with orientation difference significantly less than $\pi/2$. However, the more important consequence is computational: The Markov assumption allows us to factor the stochastic completion field into source and sink fields, each of which can be computed by a linear transformation.

To summarize, it is certainly possible to define prior distributions for boundary completion shape using devices other than a random walk (e.g., co-circularity). Unfortunately, the resulting priors will not be Markov. Consequently, the maximum likelihood completion shapes will not be curves of least energy, and it will not be possible to decompose the completion field into a product of more elementary fields.

Unlike the familiar two-dimensional isotropic random walk, where a particle's state is simply its position in the plane, the particles of the random walk described by us and Mumford (1994) possess both position and orientation. The random walk itself is defined by two elements: (1) the equations of motion and (2) a decay constant. The equations of motion describe

the change in a particle's position and orientation as a function of time; the decay constant, τ, specifies a particle's average lifetime. As part of a study of the reduction of search made possible by prior grouping in visual object recognition, Jacobs (1989) computed the probability density of the size of boundary gaps due to occlusion in random juxtapositions of a set of flat polygonal surfaces. Jacobs concluded that small gaps predominate and that incident frequency rapidly drops off with increasing size. By including a decay mechanism in the stochastic process, we are able to model the component of the completion shape probability distribution dependent on length. Because a certain fraction of particles ($1 - e^{-\frac{1}{\tau}}$) decay per unit time, longer paths are exponentially less likely. A particle's position and orientation are updated according to the following stochastic nonlinear differential equation:

$$\dot{x} = \cos\theta$$
$$\dot{y} = \sin\theta$$
$$\dot{\theta} = N(0, \sigma^2),$$

where $\dot{x}$ and $\dot{y}$ specify change in position, $\dot{\theta}$ is change in orientation (curvature), and $N(0, \sigma^2)$ is a normally distributed random variable with zero mean and variance equal to σ^2. There is a strong similiarity between the equations of motion defined here and the Kalman filter equations Cox, Rehg, and Hingorani (1993) employed in their work on edge tracking; a particle moves with constant speed in a direction that is continually changing by some random amount. The effect is that particles tend to travel in straight lines, but over time, they drift to the left or right by an amount proportional to σ^2 (for example random walks; see Figure 2). When $\sigma^2 = 0$, the motion is completely deterministic, and particles never deviate from straight paths. When $\sigma^2 = \infty$, the motion is completely random, and the stochastic process becomes a two-dimensional isotropic random walk. For this reason, we will sometimes refer to the stochastic process we define as a *diffusion* process and the parameter, σ^2 as the *diffusivity*.

4 Problem Formulation

The receptive fields of neurons in area V1 of visual cortex are retinotopically organized and narrowly tuned to stimuli of specific position and orientation (Hubel and Wiesel, 1962). Indeed, it has been suggested (see Blasdel & Obermayer, 1994; or Grossberg & Olson, 1994) that the orientation preference structure that characterizes V1 is a consequence of mapping the space of positions and orientations in the plane ($R^2 \times S^1$) onto the two-dimensional surface of the cortex (R^2) while maximizing locality. It follows that the activity of a population of neurons in V1 can be represented by a probability density function defined over $R^2 \times S^1$ and that the network of intercon-

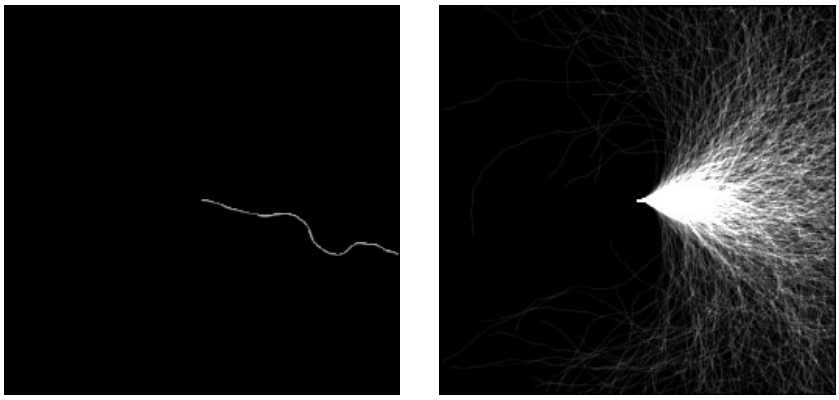

Figure 2: (Left) An example of a random walk. (Right) 1000 random walks.

nections can be represented by an order six tensor $(R^2 \times S^1) \times (R^2 \times S^1)$. Conservatively speaking, the function the network computes can be viewed as a tensor product, mapping an input probability density function to an output probability density function through a linear transformation.

Let the order six tensor G represent the transition probabilities of a Markov process defined on the three-dimensional state space, $R^2 \times S^1$, and satisfying the equations of motion defined above (G is the Green's function).[1] It follows that $G(u, v, \phi, x, y, \theta\,;\, t)$ is the probability that a random walk of length t will end in state (u, v, ϕ) given that it began in state (x, y, θ) (see Figure 3). If $p(x, y, \theta\,;\, 0)$ is the probability density function describing the position and orientation of a particle at the beginning of its random walk, then the probability that the particle will be at some position and orientation at time t is

$$p(u, v, \phi\,;\, t) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\pi}^{\pi} d\theta\; G(u, v, \phi, x, y, \theta\,;\, t) p(x, y, \theta\,;\, 0) \cdot e^{\frac{-t}{\tau}}.$$

Because the probability of a random walk of a given shape is independent of its initial position and orientation in the plane, the order six tensor, G, consists entirely of translated and rotated copies of an order three ten-

[1] Carman (1990) has emphasized the usefulness of Green's functions as descriptions of visual computations, and Carman and Welch (1992, 1993) have specifically proposed that they play a role in the computation of illusory surfaces.

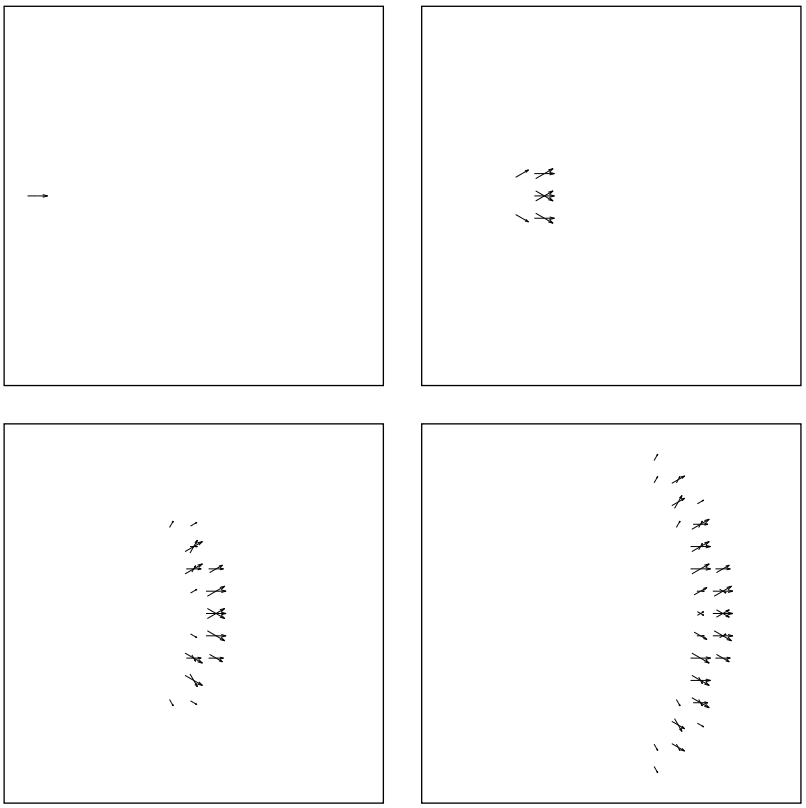

Figure 3: (Upper left) P.d.f. representing source distribution, $p(x, y, \theta; 0)$, consists of an impulse located at $(0, 0, 0)$. Although sometimes truncated for clarity, in general, the length of the arrows is proportional to the logarithm of the probability that a particle is located at that position and orientation. (Upper right) Snapshot of diffusion process at $t = 15$, that is, $p(x, y, \theta; 15)$. This is the result of convolving $G(u', v', \phi'; 15)$ with the p.d.f. representing $t = 0$. (Lower left and right) Diffusion process at $t = 31$ and $t = 47$, respectively.

sor representing the transition probabilities for random walks beginning at $(0, 0, 0)$,

$$G(u, v, \phi, x, y, \theta\,;\, t) = G(u', v', \phi', 0, 0, 0\,;\, t),$$

where (u', v', ϕ') is (u, v, ϕ) in the coordinate system defined by (x, y, θ), so that $u' = (u - x)\cos\theta + (v - y)\sin\theta$, $v' = -(u - x)\sin\theta + (v - y)\cos\theta$

and $\phi' = \phi - \theta$. Henceforward, we will write $G(u', v', \phi'; t)$ instead of $G(u', v', \phi', 0, 0, 0; t)$. The upshot is that by taking advantage of the translational and rotational symmetries of G, the tensor product can be computed by an operation similiar to convolution:

$$p(u, v, \phi;\ t) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\pi}^{\pi} d\theta\ G(u', v', \phi';\ t) p(x, y, \theta;\ 0) \cdot e^{\frac{-t}{\tau}}.$$

We will show that the stochastic completion field can be expressed as the product of a *stochastic source field* and a *stochastic sink field*. We define the stochastic source field, $p'(u, v, \phi)$, to be the fraction of paths that begin in a source state and pass through (u, v, ϕ) before they decay. If we assume that paths do not self-intersect before they decay, then the fraction of paths that pass through a given position and orientation equals the integral of the position and orientation probability density function over time:

$$p'(u, v, \phi) = \int_0^{\infty} dt\ p(u, v, \phi;\ t).$$

By changing the order of integration and defining a new Green's function, G',

$$G'(u, v, \phi) = \int_0^{\infty} dt\ G(u, v, \phi;\ t) \cdot e^{-\frac{t}{\tau}},$$

the explicit time variable, t, can be suppressed. The result is that the stochastic source field, $p'(u, v, \phi)$, can be computed by convolving the probability density function representing the source distribution with the new Green's function (see Figure 4):

$$p'(u, v, \phi) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\pi}^{\pi} d\theta\ G'(u', v', \phi')\ p(x, y, \theta;\ 0).$$

Observe that the probability[2] that a particle will pass through state (u, v, ϕ) on its way from a source state, $p \in P$, to a sink state, $q \in Q$, before it decays, is proportional to the product of (1) the probability that a particle beginning in a source will reach (u, v, ϕ) before it decays (the source field) and (2) the probability that a particle beginning at (u, v, ϕ) will reach a

[2] Strictly speaking, what we are able to compute is a relative likelihood, not a probability. This is because we do not compute the absolute probability of a particle diffusing from a source to a sink by any path at all, which would represent the normalizing constant. So although it is possible to convert these relative likelihoods to probabilities, in practice this is not necessary, because relative likelihoods suffice for the purpose of comparing competing paths.

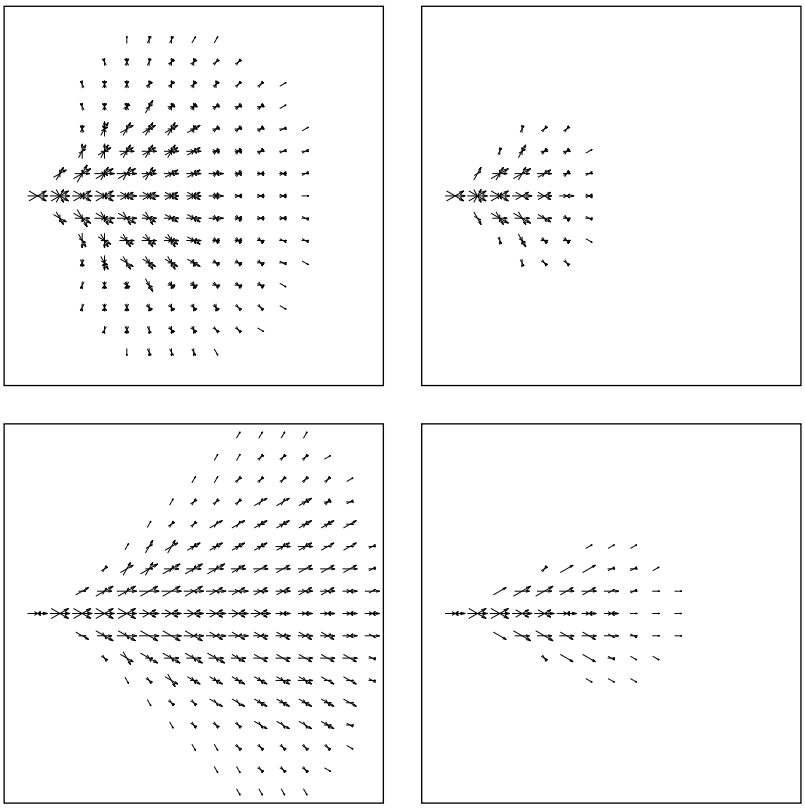

Figure 4: Stochastic source fields, $p'(u, v, \phi)$, representing the probability that a particle leaving $(0, 0, 0)$ will reach (u, v, ϕ) before it decays. Basically, these are plots of the Green's function G' (the time integral of G), for a range of diffusivities and decay constants. (Upper left) ($\sigma^2 = 0.05$, $\tau = 100$). Upper right: ($\sigma^2 = 0.05$, $\tau = 20$). (Lower left) ($\sigma^2 = 0.01$, $\tau = 100$). Lower right: ($\sigma^2 = 0.01$, $\tau = 20$).

sink before it decays (the sink field). This is an immediate consequence of the Markov property of the stochastic process. We have already described how the source field is computed. We now show that the sink field can be computed in an analogous manner.

The magnitude of the sink field at (u, v, ϕ) is the product of the probability that a particle beginning at (u, v, ϕ) will reach (x, y, θ) before it decays (i.e., $G'(x, y, \theta, u, v, \phi)$) and the probability that a sink exists at (x, y, θ) (i.e.,

$q(x, y, \theta\,;\,0)$), integrated over all (x, y, θ):

$$q'(u, v, \phi) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\pi}^{\pi} d\theta\; G'(x, y, \theta, u, v, \phi) q(x, y, \theta\,;\,0).$$

Like G, the order six tensor G' consists entirely of translated and rotated copies of an order three tensor representing the transition probabilities for random walks beginning at $(0, 0, 0)$,

$$G'(x, y, \theta, u, v, \phi) = G(x', y', \theta', 0, 0, 0),$$

where (x', y', θ') is (x, y, θ) in the coordinate system defined by (u, v, ϕ), so that $x' = (x - u)\cos\phi + (y - v)\sin\phi$, $y' = -(x - u)\sin\phi + (y - v)\cos\phi$ and $\theta' = \theta - \phi$. Consequently (like the source field), the sink field can be computed by a kind of convolution:

$$q'(u, v, \phi) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\pi}^{\pi} d\theta\; G'(x', y', \theta')\; q(x, y, \theta\,;\,0).$$

Finally, the stochastic completion field, $C(u, v, \phi)$, which represents the relative likelihood that a particle leaving a source state will pass through (u, v, ϕ) and enter a sink state before it decays, equals the product of the source and sink fields (see Figure 5):

$$C(u, v, \phi) = p'(u, v, \phi) \cdot q'(u, v, \phi).$$

At this point, it is natural to ask how the computation we have just outlined might be implemented. In particular, we wish to identify the neural loci of the representations that form the basis of our model (the source, sink, and completion fields). To begin, we note that it would be somewhat premature for us to identify the source and sink fields with a specific population in V1 because many neural populations in V1 satisfy the basic requirements of having retinotopically organized receptive fields tuned to a specific position and orientation. However, with regard to the completion field, we can be somewhat more definite and identify its locus as the neurons in V2 described by von der Heydt, Peterhans, and Baumgartner (1984). von der Heydt et al. report that the firing rate of these neurons increases when their "receptive fields" are crossed by illusory contours (of specific orientations) induced by pairs of flanking elements. Significantly, these neurons do not respond to these same elements when presented in isolation; they respond only to pairs. It was this discovery that inspired the simple feedforward neural network model proposed by Peterhans, von der Heydt, and Baumgartner (1986). At the implementation level, our model is very similar. However, while Peterhans et al. do not describe the pattern of interconnectivity between the first and second layers of their network in any detail, this pattern is defined

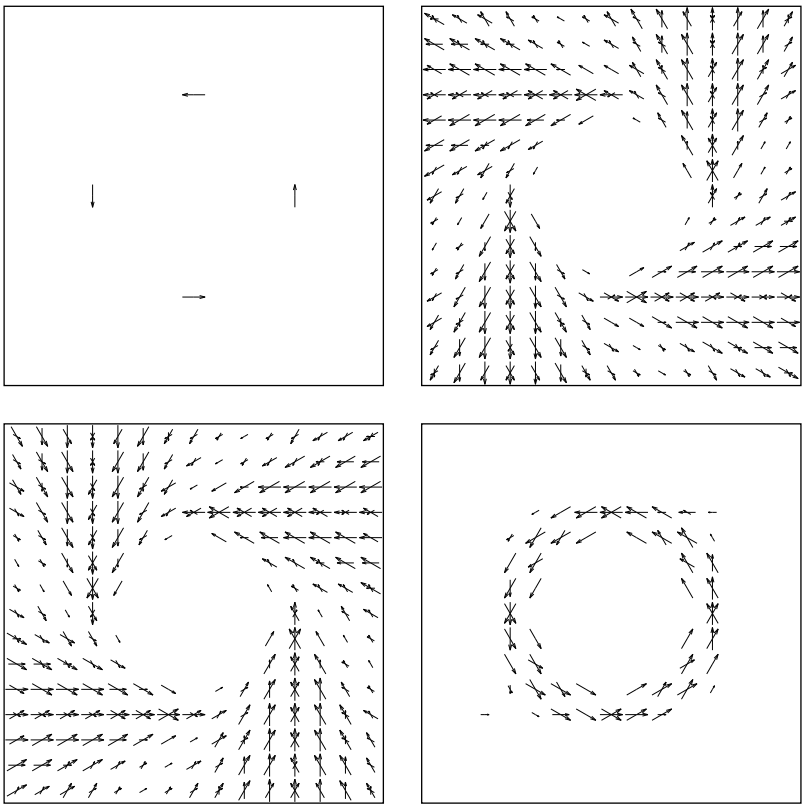

Figure 5: (Top left) Source and sink distribution $p(x, y, \theta; 0) = q(x, y, \theta; 0)$. The p.d.f. consists of four impulses equally spaced around the circumference of a circle. (Top right) The stochastic source field, $p'(u, v, \phi)$, represents the probability that a particle leaving a source will reach (u, v, ϕ) before it decays. (Bottom left) The stochastic sink field, $q'(u, v, \phi)$, represents the probability that a particle will leave state (u, v, ϕ) and enter a sink state before it decays. (Bottom right) The stochastic completion field, $C(u, v, \phi)$, is the product of source and sink fields.

in our model by the Green's function, G' (see Figure 6). Consequently, our neural network computes a stochastic completion field.

Although it is possible to solve the stochastic nonlinear differential equation directly and, in doing so, find an analytic equation for the Green's function (see Thornber & Williams, 1996), we initially solved for G' by a Monte Carlo method. All of the experimental results in this article are based

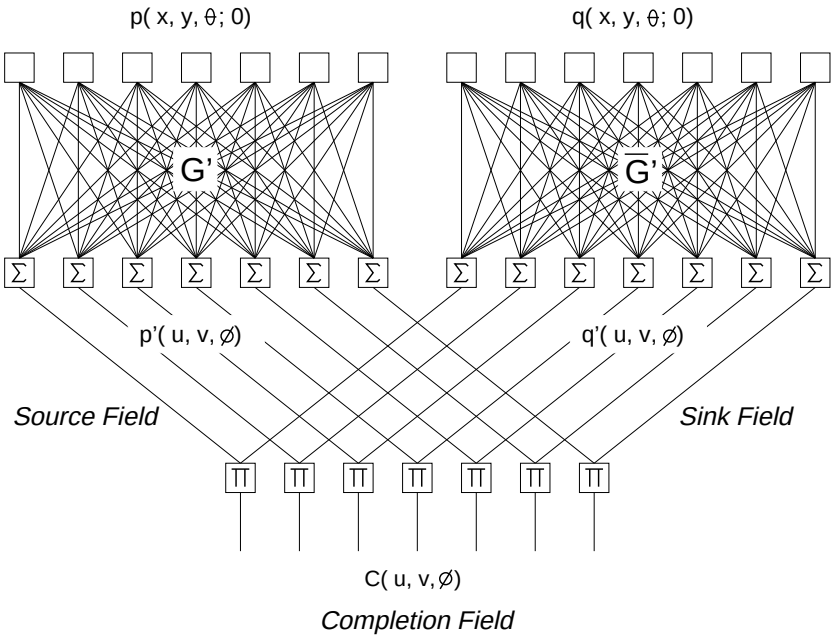

Figure 6: The stochastic completion field can be computed by a two-layer feed-forward neural network. The source field is computed by the left half of the network and the sink field by the right half. The Green's function, G', defines the network of interconnections between the first and second layers. The third layer computes the product of the source and sink fields and can be tentatively identified with the neurons in V2 described by von der Heydt, Peterhans, and Baumgartner (1984).

on an approximation of G' computed by direct simulation of the random walk for 1.0×10^6 trials on a 256×256 grid with 36 fixed orientations.[3] The particle trajectories are modeled as chains of line segments with real valued end points and with interior angles drawn from a continuous gaussian distribution. The probability that a particle beginning at $(0, 0, 0)$ will reach (u, v, ϕ) before it decays (i.e., $G'(u, v, \theta)$) is approximated by the fraction of simulated trajectories beginning at $(0, 0, 0)$ that intersect the region $(u \pm 1.0, v \pm 1.0, \phi \pm \pi/72)$. Because the end points of these segments are not confined to the integer lattice, the stochastic completion fields we compute

[3] In a computer vision application, this would be a compile-time cost, not a run-time cost.

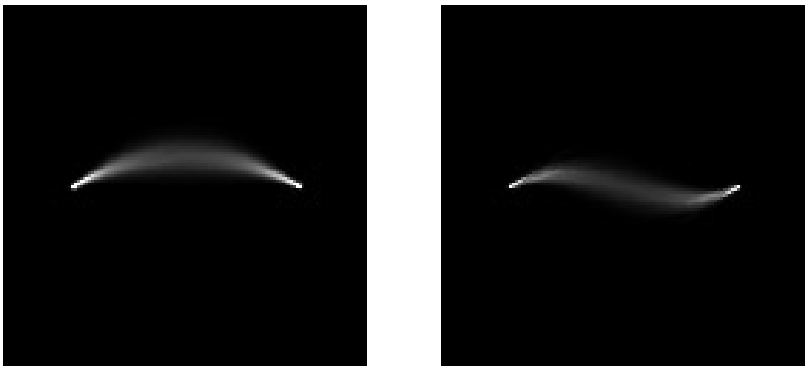

Figure 7: Example completion fields. (Left) $p_1 = (-40, 0, 30^\circ)$ and $q_1 = (40, 0, -30^\circ)$. (Right) $p_2 = (-40, 0, 30^\circ)$ and $q_2 = (40, 0, -30^\circ)$. The left source and sink pair is scaled by 1.0×10^7 and the right by 1.0×10^8.

are discrete subsamplings of continuous fields, not discrete fields computed by an approximation to the continuous process.

5 Experimental Results

As a first demonstration, let us consider two source-sink pairs. The source and sink distributions were represented by arrays of size $128 \times 128 \times 36$ and consisted of a single oriented impulse (a single nonzero value). In the first case, source and sink are positioned on a horizontal axis and are oriented symmetrically about this axis; in the second, they possess the same orientation, so that the curves of least energy joining them will contain an inflection point. The diffusivity, σ^2, equaled 0.005 and the decay constant, τ, equaled 20. Figure 7 depicts the stochastic completion field for these source and sink pairs as brightness images where brightness encodes the sum over all 36 orientations. Because, the displayed brightnesses for each pair are scaled to take maximum advantage of the limited number of gray levels (i.e., 255), it is not obvious that there is an order of magnitude difference in saliency between the first and second pair. The first source and sink is scaled by 1.0×10^7, the second by 1.0×10^8. It is sometimes assumed that illusory contours cannot contain inflection points. However, we do not believe that there is an all-or-nothing rule that prohibits inflection points (Kellman & Shipley, 1991). Rather, we believe that saliency is a continuum, and illusory contours containing inflection points might simply be an order of magnitude weaker on average.

In a second demonstration, we consider a source distribution consist-

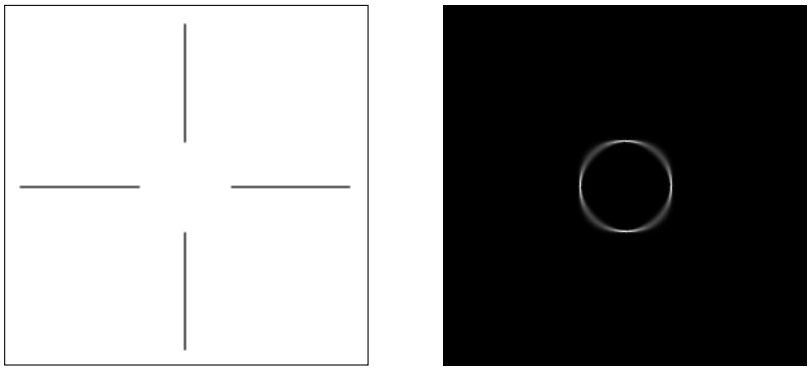

Figure 8: (Left) Ehrenstein figure. (Right) Stochastic completion field for Ehrenstein figure.

ing of four oriented impulses equally spaced around the circumference of a circle (see Figure 5, top left). This distribution is meant to represent an Ehrenstein figure (see Figure 8, left). The four impulses are located at end points of the four line segments comprising the figure and possess orientation normal to the segments. Figure 8 (right) shows the stochastic completion field, where brightness encodes the sum over all orientations. The majority of the brightness is confined to a narrow region surrounding an approximately circular ridge line. This is consistent with the shape most often reported by human observers. Interestingly, some observers report seeing a diamond shape, not a circle. We note that a diamond-shaped completion field would be produced if the diffusivity (σ^2) was very large and the particle half-life (τ) was very small.

We have also demonstrated our implementation on several well-known illusory contour figures from the visual psychology literature. Before doing this, we needed to find a principled way of translating an image into a set of sources and sinks for the diffusion process. In general, this requires a sophisticated analysis of the local image brightness structure to identify L-junctions, T-junctions, Y-junctions, and X-junctions formed by both contrast and outline edges. Classifying and measuring the multiple orientations at so-called key points (Heitger & von der Heydt, 1993) is a difficult research problem in its own right and the subject of much current research (Simoncelli & Farid, 1996; Freeman & Adelson, 1991; Michaelis & Sommer, 1994; Perona, 1992). For the moment, we use a steerable one-sided filter scheme similar to that of Simoncelli and Farid (1996) to identify orientation discontinuities (i.e., corners) formed by contrast edges and to measure the two

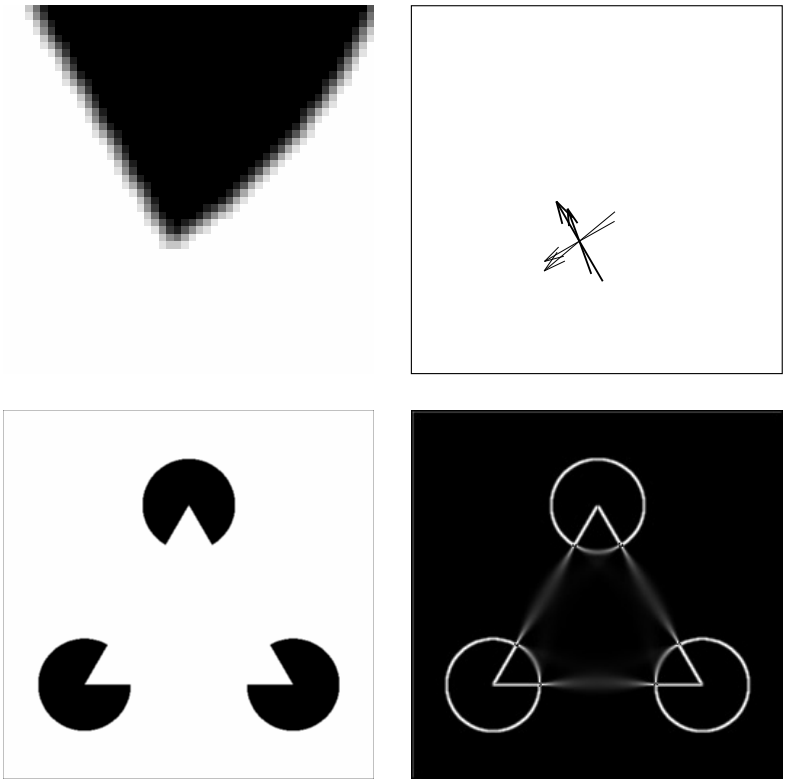

Figure 9: (Top left) A corner of a "pacman." (Top right) Sources (thin) and sinks (thick) computed by steerable one-sided filter (multiple orientations are due to anti-aliasing). (Bottom left) Kanizsa triangle. (Bottom right) Stochastic completion field summed over all orientations and superimposed on brightness gradient magnitude image. Both the illusory triangle and the three discs are completed.

orientations with precision sufficient for our purposes. The maximum and minimum of the continuous response of the steerable one-sided filter is first measured to approximately 3 degrees of accuracy. As an anti-aliasing measure, a unit mass is distributed proportionally among the two discrete orientations (10 degrees apart) straddling the nominal maximum and minimum. The maximum is interpreted as a source and the minimum as a sink (see Figure 9, top left and right). Generalizing this scheme to classify and measure the wide range of events corresponding to likely points of boundary

occlusion is the only obstacle to demonstrating our work on more realistic images.

Figure 9 (bottom left) shows the Kanizsa triangle stimulus. Key points are located at a positive maximum of curvature. Figure 9 (bottom right) shows the stochastic completion field summed over all orientations and superimposed on the brightness gradient magnitude image. Both the illusory triangle and the three discs are completed. In contrast with the results of Heitger and von der Heydt (1993), no nonmaximum suppression needs to be employed.

Figure 10 (top) shows the Kanizsa "paisley" stimulus. Key points are located at negative minima of curvature. Figure 10 (bottom) shows the stochastic completion field summed over all orientations and superimposed on the brightness gradient magnitude image.

Figure 11 (top) shows a complex illusory contour figure designed by Kanizsa (1979). Key points are located at positive maxima of curvature. It is useful to compare the stochastic completion field computed for this display to the middle panel of Figure 1. Because our diffusion process has no knowledge of the topology of surfaces, the stochastic completion field, shown in Figure 11 (bottom left), contains potential completions that are not perceived by human subjects. Potential completions required to complete the four rectangles are among the most salient, however. Figure 11 (bottom right) shows the logarithm of the stochastic completion field summed over all orientations. In the logarithm image, many additional completions of significantly lower average likelihood become visible. Included among these are those required to complete the four black discs and eight black squares perceived by human subjects.

6 Conclusion

It is widely acknowledged that perceptual organization (segmentation, grouping) is among the most difficult problems facing researchers in computer vision today. Current structure from motion algorithms produces estimates of depth only for isolated points, yet true surface representations are required to compute a stable grasp or to plan an unobstructed path through free space. Current object recognition systems do not work on large model bases, nor do they function robustly in the presence of occlusion. Our work advances the state of the art in perceptual organization by providing a plausible model of illusory contour formation based on a diffusion process. We have shown how curves of least energy interpolating a set of edge fragments can be computed using biologically plausible algorithms and representations. Significantly, this is accomplished without using numerical relaxation or other explicit minimization but instead relies on the fact that the probability that a particle following a random walk will pass through a particular position and orientation on a path joining two image

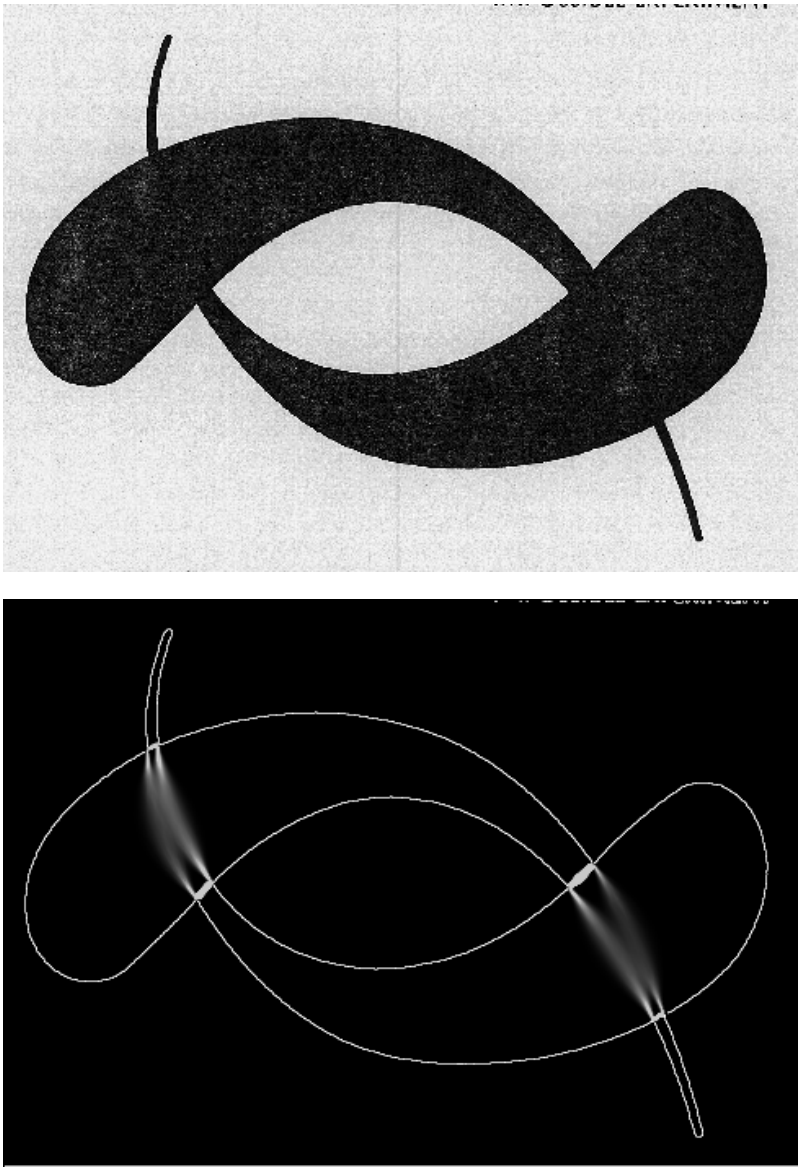

Figure 10: (Top) Kanizsa's "paisleys." (Bottom) Stochastic completion field integrated over all orientations and superimposed on brightness gradient magnitude image. It is interesting to note that the average likelihoods of the shorter completions (perceived modally) are several orders of magnitude greater than the average likelihoods of the longer completions (perceived amodally).

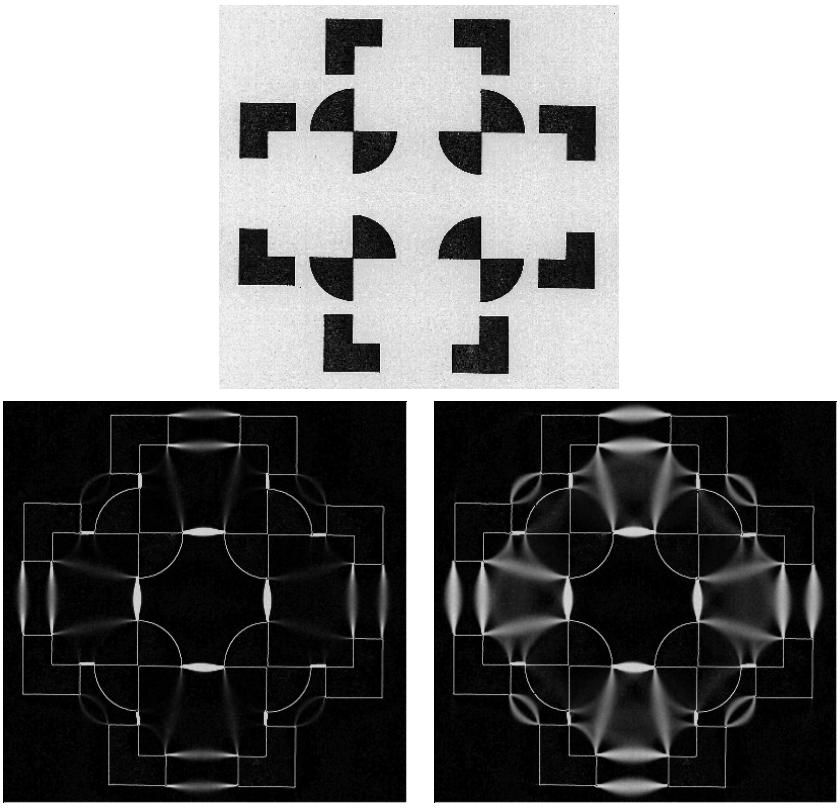

Figure 11: (Top) A complex figure (Kanizsa, 1979). (Bottom) Stochastic completion field integrated over all orientations and superimposed on brightness gradient magnitude image. Logarithm of stochastic completion field integrated over all orientations.

measurements can be computed directly as the product of two vector-field convolutions.

The most related work is by Heitger and von der Heydt (1993), Guy and Medioni (1993) and Shashua and Ullman (1988). Although we owe an intellectual debt to these three articles, none proceeds cleanly from a definition of what they want to compute (a function) to a method of computing it (an algorithm). Our main contribution is to show how outwardly similiar algorithms and representations follow immediately from a clearly stated assumption: that the prior probability distribution of boundary completion shapes can be modeled by a random walk.

Appendix

In this appendix, we demonstrate the relationship between the diffusion process we have described and so-called curves of least energy. To accomplish this, we employ a discrete-time approximation to the continuous equations of motion. For the discrete-time random walk, we show that the most likely path between a source and a sink is the curve of least energy connecting them. This tells us that, in the limit, as the time-step size becomes small, the most likely path between a source and a sink converges to the continuous curve of least energy. From this, we conclude that our model of diffusion concisely encodes the prior assumption that the most likely shape of an occluded section of an object's boundary is the curve of least energy.

Suppose that a particle begins its random walk at a source p and follows a trajectory Γ, which consists of n unit length steps, along with n changes in angle denoted by $\kappa_1, \ldots, \kappa_n$, ending at sink q. That is, its trajectory is an n-sided polygonal arc, comprising unit length segments and with exterior angles denoted by κ_i. From the definition of the random walk, we know that the density function on the set of paths that such a particle may take is given by

$$g(\Gamma_{pq}) = \frac{1}{\int_{\Gamma_q} f(\Gamma_p) d\Gamma_p} \prod_{i=1}^{n} e^{-\frac{1}{\tau}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{k_i^2}{2\sigma^2}},$$

where $\int_{\Gamma_q} f(\Gamma_p) d\Gamma_p$ indicates integration of the density function on the set of paths beginning in source p (i.e., $f(\Gamma_p)$) over all paths ending in sink q (see Ash, 1972, for a discussion of conditional probabilities on continuous functions). Taking the logarithm of both sides (and because $\int_{\Gamma_q} f(\Gamma_p) d\Gamma_p$ is constant for given p and q),

$$\log(g(\Gamma_{pq})) + C = n\left(-\frac{1}{\tau} - \log(\sigma\sqrt{2\pi})\right) - \sum_{i=1}^{n} \frac{k_i^2}{2\sigma^2}. \tag{A.1}$$

We now relate this expression to the energy of a continuous curve. This energy is defined as

$$\alpha \int_{\Gamma} \kappa(t)^2 dt + \beta \int_{\Gamma} dt,$$

where $\kappa(t)$ is the curvature at $\Gamma(t)$ and α and β are constants that weight the cost for the length relative to the cost for the total squared curvature. Although the energy of any curve with a curvature discontinuity is infinite, it is natural to consider an analogous quantity for polygonal arcs based on number of segments (n); and sum of the squared exterior angles ($\sum_{i=1}^{n} k_i^2$). This is precisely the negative of the right side of equation (A.1), with $\alpha = 1/2\sigma^2$ and

$\beta = 1/\tau + \log(\sigma\sqrt{2\pi})$. This shows that the log likelihood that a random walk will follow any given polygonal path is linearly related to the energy of that path for a natural discrete formulation of energy. Therefore, the maximum likelihood polygonal path between two points is the discrete curve of least energy. Because the discrete-time random walk converges to a continuous-time random walk as the time-step size, velocity, and diffusivity approach zero, it follows that the continuous equations of motion concisely encode our prior assumption concerning the probability distribution of completion shape.

Acknowledgments

We gratefully acknowledge many helpful conversations with Karvel Thornber. Thanks also to Ingemar Cox and Bill Bialek.

References

Alter, T., & Basri, R. (1996). Extracting salient contours from images: An analysis of the saliency network, *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, San Francisco, pp. 13–20.

Ash, R. (1972). *Real analysis and probability*. San Diego: Academic Press.

Blasdel, G., & Obermeyer, K. (1994). Putative strategies of scene segmentation in monkey visual cortex. *Neural Networks*, *7*(6–7), 865–881.

Carman, G. J. (1990). *Mappings of the cerebral cortex*. Unpublished doctoral dissertation, California Institute of Technology, Pasadena, CA.

Carman, G. J., & Welch, L. (1992). Three-dimensional illusory contours and surfaces. *Nature*, *360*, 585–587.

Carman, G. J., & Welch, L. (1993). Position, orientation and shape of real and illusory three-dimensional surfaces. *Investigative Opthalmology and Visual Science (ARVO)*, *34*(4), 1131.

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.

David, C., & Zucker, S. W. (1990). Potentials, valleys and dynamic global coverings. *Inernational Journal of Computer Vision*, *5*(3), 219–238.

Freeman, W. T., & Adelson, E. H. (1991). The design and use of steerable filters for image analysis, enhancement and wavelet representation. *IEEE Pattern Analysis and Machine Intelligence*, *13*, 891–906.

Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception: Boundary completion, and illusory figures, and neon color spreading. *Psychological Review*, *92*, 173–211.

Grossberg, S., & Olson, S. J. (1994). Rules for the cortical map of ocular dominance and orientation columns. *Neural Networks*, *7*(6–7), 883–894.

Guy, G., & Medioni, G. (1993). Inferring global perceptual contours from local features. *Inernational Journal of Computer Vision*, *20*, 113–133.

Heitger, R., & von der Heydt, R. (1993). A computational model of neural contour processing, figure-ground and illusory contours. *Proc. of 4th Intl. Conf. on Computer Vision*, Berlin, Germany.

Horn, B. K. P. (1981). The curve of least energy (MIT AI Lab Memo No. 612). Cambridge, MA: MIT.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, *160*, 106–154.

Huffman, D. A. (1971). Impossible objects as nonsense sentences. *Machine Intelligence*, *6*, 295–323.

Jacobs, D. (1989). Grouping for recognition (MIT AI Lab Memo No. 1177). Cambridge, MA: MIT.

Kanizsa, G. (1979). *Organization in vision*. New York: Praeger.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141–221.

Michaelis, M., & Sommer, G. (1994). Junction classification by multiple orientation detection. *Proc. of European Conf. on Computer Vision*, pp. 101–108.

Mumford, D. (1994). Elastica and computer vision. *Algebraic Geometry and Its Applications*. Chandrajit Bajaj (Ed.). New York: Springer-Verlag.

Parent, P., & Zucker, S. W. (1989). Trace inference, curvature consistency and curve detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*, pp. 823–889.

Perona, P. (1992). Steerable, scalable kernels for edge detection and junction analysis. *Proc. of European Conf. on Computer Vision*, pp. 3–18.

Peterhans, E., von der Heydt, R., & Baumgartner, G. (1986). Neuronal responses to illusory contour stimuli reveal stages of visual cortical processing. In J. Pettigrew, K. Sanderson, & W. Levick, (Eds.), *Visual Neuroscience*. Cambridge: Cambridge University Press.

Shashua, A., & Ullman, S. (1988). Structural saliency: The detection of globally salient structures using a locally connected network. *2nd Intl. Conf. on Computer Vision*, Clearwater, FL, pp. 321–327.

Simoncelli, E. P., & Farid, H. (1996). Steerable wedge filters for local orientation analysis. *IEEE Transactions on Image Processing*, *5*(9), 1377–1382.

Thornber, K. K., & Williams, L. R. (1996). Analytic solution of stochastic completion fields. *Biological Cybernetics*, *75*, 141–151.

Ullman, S. (1976). Filling in the gaps: The shape of subjective contours and a model for their generation. *Biological Cybernetics*, *21*, 1–6.

von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, *224*, 1260–1262.

Williams, L. R., & Hanson, A. R. (1996). Perceptual completion of occluded surfaces. *Computer Vision and Image Understanding*, *64*(1), 1–20.

Received May 5, 1995; accepted July 12, 1996.