Euclidean Group Invariant Computation of Stochastic Completion Fields Using Shiftable-Twistable Functions

John W. Zweck and Lance R. Williams

Dept. of Computer Science
University of New Mexico
Albuquerque, NM 87131, USA.

Abstract. We describe a method for computing the likelihood that a completion joining two contour fragments passes through any given position and orientation in the image plane, that is, a method for completing the boundaries of partially occluded objects. Like computations in primary visual cortex (and unlike all previous models of contour completion in the human visual system), our computation is Euclidean invariant. This invariance is achieved in a biologically plausible manner by representing the input, output, and intermediate states of the computation in a basis of shiftable-twistable functions. The spatial components of these functions resemble the receptive fields of simple cells in primary visual cortex. Shiftable-twistable functions on the space of positions and directions are a generalization of shiftable-steerable functions on the plane.

1 Introduction

Any computational model of human visual information processing must reconcile two apparently contradictory observations. First, computations in primary visual cortex are largely *Euclidean invariant*—an arbitrary rotation and translation of the input pattern of light falling on the retina produces an identical rotation and translation of the output of the computation. Second, simple calculations based on the size of primary visual cortex (60 mm $\times$ 80 mm) and the observed density of cortical hypercolumns (4/mm^2) suggest that the discrete spatial sampling of the visual field is exceedingly sparse [24]. The apparent contradiction becomes clear when we ask the following questions: How is this remarkable invariance achieved in computations performed by populations of cortical neurons with broadly tuned receptive fields centered at so few locations? Why doesn't our perception of the world change dramatically when we tilt our head by 5 degrees?[1]

[1] Ulf Eyesel asks a related question in a recent *Nature* paper [5]:

> *"On average, a region of just 1 mm^2 on the surface of the cortex will contain all possible orientation preferences, and, accordingly, can analyze orientation for one small area of the visual field. This topographical arrangement allows closely spaced objects with different orientations to interact. But it also means*

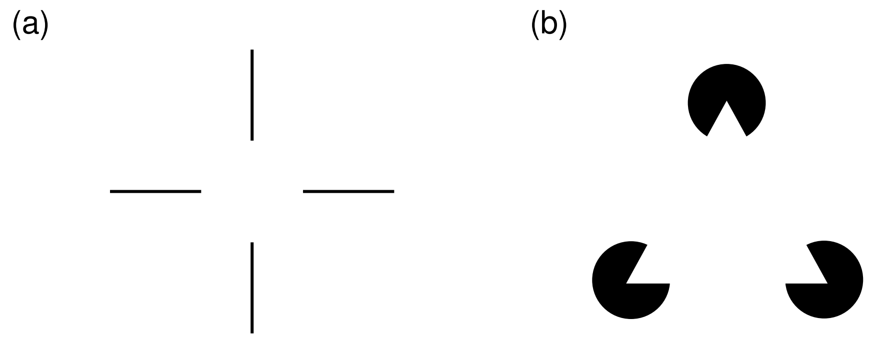

Fig. 1. (a) Ehrenstein Figure **(b)** Kanizsa Triangle

One of the main goals of our research is to show how the sparse and seemingly haphazard nature of the sampling of the visual field can be reconciled with the Euclidean invariance of visual computations. To realize this goal, we introduce the notion of a shiftable-twistable basis of functions on the space, $\mathbf{R}^2 \times S^1$, of positions and directions. This notion is a generalization of the notion of a shiftable-steerable basis of functions on the plane, $\mathbf{R}^2$, introduced by Freeman, Adelson, Simoncelli, and Heeger in two seminal papers [6, 18]. Freeman and Adelson [6] clearly appreciated the importance of the issues raised above when they devised the notion of a steerable basis to implement rotationally invariant computations. In fact, for computations in the plane the contradictions discussed above were largely resolved with the introduction by Simoncelli *et al.* [18] of the shiftable-steerable pyramid transform, which was specifically designed to perform Euclidean invariant computations on $\mathbf{R}^2$. The basis functions in the shiftable-steerable pyramid are very similar to simple cell receptive fields in primary visual cortex. However, many computations in V1 and V2 likely operate on functions of the space of positions and directions, $\mathbf{R}^2 \times S^1$, rather than on functions of the plane, $\mathbf{R}^2$ (e.g., [8, 9, 13, 16, 17, 21, 22, 25]). Consequently, we propose that shiftability-twistability (in addition to shiftability-steerability) is the property which binds sparsely distributed receptive fields together functionally to perform Euclidean invariant computations in visual cortex.

In this article, we describe a new algorithm for completing the boundaries of partially occluded objects. This algorithm is based on a computational theory of contour completion in primary and secondary visual cortex developed in recent years by Williams and colleagues [19–22]. Like computations in V1 and V2, and unlike previous models of illusory contour formation in the human visual system, our computation is Euclidean invariant. This invariance is achieved by

---

that a continuous line across the whole visual field would be cortically depicted in a patchy, discontinuous fashion. How can the spatially separated elements be bound together functionally?"

representing the input, output, and intermediate states of the computation in a basis of shiftable-twistable functions.

Mumford [15] proposed that the probability distribution of natural shapes can be modeled by particles traveling with constant speed in directions given by Brownian motions. More recently, Williams and Jacobs [21] defined the *stochastic completion field* to be the distribution of particle trajectories joining pairs of position and direction constraints, and showed how it could be computed in a neural network.

The neural network described in [22] is based on Mumford's observation that the evolution in time of the probability density function (p.d.f.) representing the position, (x, y), and direction, θ, of the particle can be modeled as a set of independent advection equations acting in the (x, y) dimension coupled in the θ dimension by the diffusion equation [15]. Unfortunately, solutions of this *Fokker-Planck* equation computed by numerical integration on a rectangular grid do not exhibit the robust invariance under rotations and translations which characterizes the output of computations performed in primary visual cortex. Nor does any other existing model of contour completion, sharpening, or saliency (e.g., [8, 9, 13, 16, 17, 21, 22, 25]).

Our new algorithm computes stochastic completion fields in a Euclidean invariant manner. Figure 2 (left) is a picture of the stochastic completion field due to the Kanizsa Triangle stimulus in Figure 1(b). Figure 2 (right) shows the stochastic completion field due to a rotation and translation of the (input) Kanizsa Triangle. The Euclidean invariance of our algorithm can be seen by observing that the (output) stochastic completion field on the right in Figure 2 is itself a rotation and translation of the stochastic completion field on the left, by the same amount.

2 Relevant Neuroscience

Our new Euclidean invariant algorithm was motivated, in part, by the following experimental findings. To begin with, the receptive fields of simple cells, which have been traditionally described as edge (or bar) detectors, can be accurately modeled using two-dimensional Gabor functions [3, 14], which are the product of a Gaussian (localized in position) and a harmonic grating (localized in orientation and spatial frequency). Gabor functions are well suited to the purpose of encoding visual information, since, by the Heisenberg Uncertainty Principle, they are the unique functions which are maximally localized in both space and frequency.

The sampling of the visual field in V1 is quite sparse—there are about about 100×100 hypercolumns, with receptive fields of about 5 scales and 16 orientations in each hypercolumn. Neglecting size (and phase), a simple cell receptive field can be parameterized by its position and orientation. The spatial distribution of these two parameters, known as *orientation preference structure*, is an attempt (on the part of evolution) to smoothly map the three-dimensional parameter space, $\mathbf{R}^2 \times S^1$, of edge positions and orientations onto the two-dimensional surface, of the

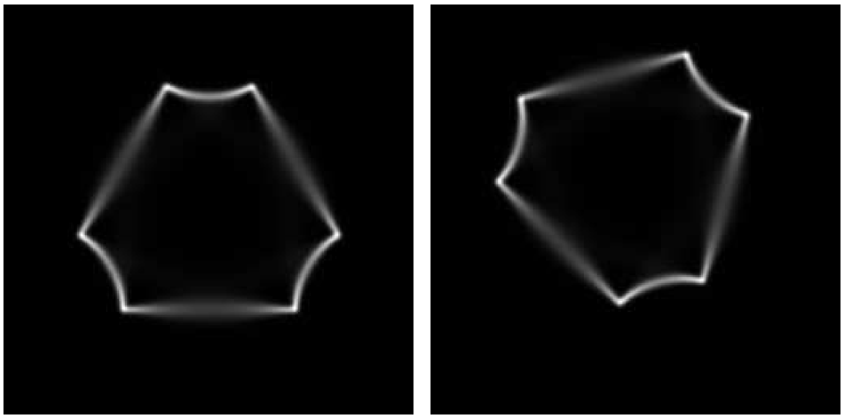

Fig. 2. Stochastic completion fields: Of Kanizsa Triangle (*left*) and after the initial conditions have been rotated and translated (*right*)

visual cortex, $\mathbf{R}^2$. Due to the differences in dimensionality, orientation preference structure is punctuated by so-called *pinwheels*, which are the singularities in this mapping [1].

As a first approximation, a neuron's response to an arbitrary grey-level image can be modeled as the L^2-inner product of the image with the neuron's receptive field. These experimental observations suggested to Daugman [4] that an ensemble of simple cell receptive fields can be regarded as performing a wavelet transform of the image, in which the responses of the neurons correspond to the transform coefficients and the receptive fields correspond to the basis functions.

Recent experiments have demonstrated that the response of simple cells in V1 can be modulated by stimuli outside the classical receptive fields. Apart from underscoring the limitations of the classical (linear) model, they suggest a function for the long-range connections which have been observed between simple cells. For example, in a recent experiment, Gilbert [7] has demonstrated that a short horizontal bar stimulus can modulate the response of simple cells whose receptive fields are located at a significant horizontal distances from the bar, and which have a similar orientation preference to the bar. Non-linear long-range effects have also been observed in secondary visual cortex. For example, von der Heydt *et al.*[10] reported that the firing rate of certain neurons in V2 increases when their "receptive fields" are crossed by illusory contours (of specific orientations) which are induced by pairs of bars flanking the receptive field. Significantly, these neurons do not respond to these same bars presented in isolation—they only respond to pairs.[2]

[2] These experiments suggests that the source and sink fields, which are intermediate representations in Williams and Jacobs model of illusory contour formation, could be

Although our new contour completion algorithm does not provide a model for illusory contour formation in the brain which is realistic in every respect, it does have several features which are biologically plausible, none of which are found in previous algorithms, e.g., [8, 9, 13, 16, 17, 21, 22, 25]. These features are that (1) all states of the computation be represented in a wavelet-like basis of functions which are localized in both space and frequency (spatial localization allows the computation to be performed in parallel); (2) the computation operates on the coefficients in the wavelet-like transform and can be implemented in a neural network; (3) the computation is Euclidean invariant; and (4) it is accomplished using basis functions with centers lying on a (relatively) sparse grid in the image plane.

3 Shiftable-Twistable Bases

Many visual and image processing tasks are most naturally formulated in the continuum and are invariant under a group of symmetries of the continuum. The Euclidean group, of rotations and translations, is one example of a continuous symmetry group. However, because discrete lattices are not preserved by the action of continuous symmetry groups, the natural invariance of a computation can be easily lost when it is performed in a discrete network. In this section we will introduce the notion of a shiftable-twistable basis and show how it can be used to implement discrete computations on the continuous space of positions and directions in a way which preserves their natural invariance.

In image processing, the input and output are functions on $\mathbf{R}^2$, and the appropriate notion of the invariance of computations is *Euclidean invariance*—any rotation and translation of the input should produce an identical rotation and translation of the output. Simoncelli *et al.* [6, 18] introduced the notion of a shiftable-steerable basis of functions on $\mathbf{R}^2$, and showed how it can be used to achieve Euclidean invariance in discrete computations for image enhancement, stereo disparity measurement, and scale-space analysis.

Given the nature of simple cell receptive fields, the input and output of computations in primary visual cortex are more naturally thought of as functions defined on the continuous space, $\mathbf{R}^2 \times S^1$, of positions, $\boldsymbol{x} = (x, y)$, in the plane, $\mathbf{R}^2$, and directions, θ, in the circle, S^1. For such computations the appropriate notion of invariance is determined by those symmetries, $T_{\boldsymbol{x}_0,\theta_0}$, of $\mathbf{R}^2 \times S^1$, which perform a shift in $\mathbf{R}^2$ by $\boldsymbol{x}_0$, followed by a twist in $\mathbf{R}^2 \times S^1$ through an angle, θ_0. A *twist* through an angle, θ_0, consists of two parts: (1) a rotation, R_{θ_0}, of $\mathbf{R}^2$ and (2) a translation in S^1, both by θ_0. The symmetry, $T_{\boldsymbol{x}_0,\theta_0}$, which is called a *shift-twist transformation*[3], is given by the formula,

$$T_{(\boldsymbol{x}_0,\theta_0)}(\boldsymbol{x},\theta) \;=\; (R_{\theta_0}(\boldsymbol{x}-\boldsymbol{x}_0)\,,\,\theta-\theta_0)\ . \tag{3.1}$$

represented by populations of simple cells in V1, and that the stochastic completion field, which is the product of the source and sink fields, could be represented in V2.

[3] The relationship between shift-twist transformations and computations in V1 was described by Williams and Jacobs in [21] and more recently by Kalitzin *et al.* [12] and Cowan [2].

A visual computation on $\mathbf{R}^2 \times S^1$ is called *shift-twist invariant* if, for all $(\boldsymbol{x}_0, \theta_0) \in \mathbf{R}^2 \times S^1$, a shift-twist of the input by $(\boldsymbol{x}_0, \theta_0)$ produces an identical shift-twist of the output.

Correspondingly, we define a *shiftable-twistable basis*[4] of functions on $\mathbf{R}^2 \times S^1$ to be a set of functions on $\mathbf{R}^2 \times S^1$ with the property that whenever a function, $P(\boldsymbol{x}, \theta)$, is in their span, then so is $P(T_{\boldsymbol{x}_0,\theta_0}(\boldsymbol{x}, \theta))$, for every choice of $(\boldsymbol{x}_0, \theta_0)$ in $\mathbf{R}^2 \times S^1$. As such, the notion of a shiftable-twistable basis on $\mathbf{R}^2 \times S^1$ generalizes that of a shiftable-steerable basis on $\mathbf{R}^2$.

Shiftable-twistable bases can be constructed as follows. First we recall Simoncelli's concept of the shiftability of a function, which is closely related to the Shannon-Whittaker Sampling Theorem. A periodic function, $\psi(x)$, of period X, is *shiftable* if there is an integer, K, such that the shift of ψ by an arbitrary amount, x_0, can be expressed as a linear combination of K basic shifts of ψ, i.e., if there exist *interpolation functions*, $b_k(x_0)$, such that

$$\psi(x - x_0) \;=\; \textstyle\sum_{k=0}^{K-1} \; b_k(x_0) \; \psi(x - k\Delta) \; , \tag{3.2}$$

where $\Delta = X/K$ is the basic shift amount. The simplest shiftable function in one dimension is a pure harmonic signal, $e^{i\omega x}$, in which case $K = 1$. More generally, Simoncelli *et al.* [18] proved that any band-limited function is shiftable. In fact, if the set of non-zero Fourier series frequencies of ψ is (a subset of) $B = \{\omega_0, \omega_0+1, \ldots, \omega_0+K-1\}$, then ψ can be shifted using the K interpolation functions, $b_k(x_0) = b(x_0 - k\Delta)$, where $b(x)$ is the complex conjugate of the perfect bandpass filter constructed from the set of K frequencies, B. In particular, note that the interpolation functions only depend on the set of non-zero frequencies of ψ, and not on ψ itself.

Strictly speaking, since they are not band-limited, functions such as Gabors are not shiftable. Nevertheless, for all intents and purposes, they can be shifted by choosing the set, B, to consist of all Fourier series frequencies, ω, of ψ, such that the Fourier amplitude, $|\hat{\psi}(\omega)|$, is *essentially non-zero* (i.e., it exceeds some small threshold value). Such functions will be called *effectively shiftable.*

Let $\Psi(\boldsymbol{x}, \theta)$ be a function on $\mathbf{R}^2 \times S^1$ which is periodic (with period X) in both spatial variables, $\boldsymbol{x}$. In analogy with the definition of a shiftable-steerable function on $\mathbf{R}^2$, we say that Ψ is *shiftable-twistable* on $\mathbf{R}^2 \times S^1$ if there are integers, K and M, and interpolation functions, $b_{\boldsymbol{k},m}(\boldsymbol{x}_0, \theta_0)$, such that, for each $(\boldsymbol{x}_0, \theta_0) \in \mathbf{R}^2 \times S^1$, the shift-twist of Ψ by $(\boldsymbol{x}_0, \theta_0)$ is a linear combination of a finite number of basic shift-twists of Ψ by amounts $(\boldsymbol{k}\Delta, m\Delta_\theta)$, i.e., if

$$\Psi(T_{\boldsymbol{x}_0,\theta_0}(\boldsymbol{x}, \theta)) \;=\; \textstyle\sum_{\boldsymbol{k},m} \; b_{\boldsymbol{k},m}(\boldsymbol{x}_0, \theta_0) \; \Psi(T_{\boldsymbol{k}\Delta, m\Delta_\theta}(\boldsymbol{x}, \theta)) \; . \tag{3.3}$$

Here $\Delta = X/K$ is the *basic shift amount* and $\Delta_\theta = 2\pi/M$ is the *basic twist amount.* The sum in equation (3.3) is taken over all pairs of integers, $\boldsymbol{k} = (k_x, k_y)$, in the range, $0 \le k_x, k_y < K$, and all integers, m, in the range, $0 \le m < M$. For many shiftable-twistable bases, the interpolation functions, $b_{\boldsymbol{k},m}(\boldsymbol{x}_0, \theta_0)$, are defined in terms of Simoncelli's one-dimensional interpolation functions, $b_k(x_0)$.

[4] We use this terminology even though the basis functions need not be linearly independent.

The simplest example of a shiftable-steerable basis is the Gaussian-Fourier basis, $G_{\boldsymbol{k},\omega}$, which is the product of a shiftable-steerable basis of Gaussians in $\boldsymbol{x}$ and a Fourier series basis in θ. Let $g(\boldsymbol{x}) = \frac{1}{\nu}e^{-\|\boldsymbol{x}\|^2/2\nu^2}$ be a radial Gaussian of standard deviation, ν. We regard g as a periodic function of period, X, which is chosen to be much larger than ν, so that g defines a smooth periodic function. For each frequency, ω, we define $G_\omega(\boldsymbol{x},\theta) = g(\boldsymbol{x})e^{i\omega\theta}$. Also, given a choice of a shift amount, Δ, so that $K = X/\Delta$ is an integer, we define the *Gaussian-Fourier basis functions*, $G_{\boldsymbol{k},\omega}$, by

$$G_{\boldsymbol{k},\omega}(\boldsymbol{x},\theta) \;=\; g(\boldsymbol{x}-\boldsymbol{k}\Delta)\; e^{i\omega\theta} \; . \tag{3.4}$$

The following proposition implies that the Gaussian-Fourier basis is shiftable-twistable.

Proposition 1. *The periodic function, G_ω, (of period X) is effectively shiftable-twistable. More precisely, let $M = 1$ and let K be the number of essentially non-zero Fourier series coefficients of the factor, $g_X(x) = e^{-x^2/2\nu^2}$, of $g(\boldsymbol{x})$. Then,*

$$G_\omega(T_{\boldsymbol{x}_0,\theta_0}(\boldsymbol{x},\theta)) \;=\; \sum_{\boldsymbol{k}} b_{\boldsymbol{k},\omega}(\boldsymbol{x}_0,\theta_0)\; G_{\boldsymbol{k},\omega}(\boldsymbol{x},\theta) \; , \tag{3.5}$$

where the interpolation functions are given by

$$b_{\boldsymbol{k},\omega}(\boldsymbol{x}_0,\theta_0) \;=\; e^{-i\omega\theta_0}\; b_{\boldsymbol{k}}(\boldsymbol{x}_0) \; . \tag{3.6}$$

Here $b_{\boldsymbol{k}}(\boldsymbol{x}_0) = b(\boldsymbol{x}_0 - \boldsymbol{k}\Delta)$, where $b(\boldsymbol{x}_0)$ is the complex conjugate of the perfect bandpass filter constructed from the set of K^2 essentially non-zero Fourier series coefficients, $\boldsymbol{\eta}$, of $g(\boldsymbol{x})$.

For certain computations, the input can be easily represented in a Gaussian-Fourier basis. For example, suppose that the input is modeled as a linear combination of fine scale three-dimensional Gaussians, centered at arbitrary points, $(\boldsymbol{x}_0,\theta_0)$, in $\mathbf{R}^2 \times S^1$. Since the input is the product of a Gaussian in $\boldsymbol{x}$ and a Gaussian in θ it can be represented in a *single scale* Gaussian-Fourier basis as follows. First, the Gaussian in θ is represented in the Fourier basis using the standard analysis and synthesis formulae for Fourier series. Second, if the two-dimensional input Gaussians in $\boldsymbol{x}$ are chosen to be shifts of the basis function, $g(\boldsymbol{x})$, then we can use Proposition 1 to represent the input Gaussians in $\boldsymbol{x}$ in the Gaussian basis.

A somewhat more biologically plausible basis is the *complex directional derivative of Gaussian (CDDG)-Fourier basis*, which is very similar to the previous example, except that the Gaussian, $g(\boldsymbol{x})$, is replaced by its complex directional derivative in the direction of the complex valued vector, $[1,i]^{\mathrm{T}}$. The CDDG looks more like the receptive field of a simple cell in V1 than a Gaussian does. Also the CDDG is a wavelet, whereas the Gaussian is not.

4 Stochastic Completion Fields

In their computational theory of illusory contour formation, Williams and Jacobs [21] argued that, given a prior probability distribution of possible completion shapes, the visual system computes the local image plane statistics of the distribution of all possible completions, rather than simply the most probable completion. This view is in accord with human experience—some illusory contours are more salient than others, and some appear sharper than others. They defined the notion of a stochastic completion field to model illusory contours in a probabilistic manner. The stochastic completion field is a probability density function (p.d.f.) on the space, $\mathbf{R}^2 \times S^1$, of positions, $\boldsymbol{x} = (x, y)$, in the plane, $\mathbf{R}^2$, and directions, θ, in the circle, S^1. It is defined in terms of a set of position and direction constraints representing the beginning and ending points of a set of contour fragments (called *sources* and *sinks*), and a prior probability distribution of completion shapes, which is modeled as the set of paths followed by particles traveling with constant speed in directions described by Brownian motions [15]. The magnitude of the *stochastic completion field*, $C(\boldsymbol{x}, \theta)$, is the probability that a completion from the prior probability distribution will pass through $(\boldsymbol{x}, \theta)$ on a path joining two of the contour fragments. Williams and Jacobs [21] showed that the stochastic completion field could be factored into a *source field* and a *sink field*. The source field, $P'(\boldsymbol{x}, \theta)$, represents the probability that a contour beginning at a source will pass through $(\boldsymbol{x}, \theta)$ and the sink field, $Q'(\boldsymbol{x}, \theta)$, represents the probability that a contour beginning at $(\boldsymbol{x}, \theta)$ will reach a sink. The completion field is

$$C(\boldsymbol{x}, \theta) \;=\; P'(\boldsymbol{x}, \theta) \cdot \; Q'(\boldsymbol{x}, \theta) \; . \tag{4.1}$$

The source (or sink) field itself is obtained by integrating a probability density function, $P(\boldsymbol{x}, \theta\,; t)$, over all positive times, t, where $P(\boldsymbol{x}, \theta\,; t)$ represents the probability that a particle beginning at a source reaches $(\boldsymbol{x}, \theta)$ at time t,

$$P'(\boldsymbol{x}, \theta) \;=\; \int_0^\infty P(\boldsymbol{x}, \theta\,; t) dt \; . \tag{4.2}$$

Mumford [15] observed that P evolves according to a Fokker-Planck equation of the form,

$$\frac{\partial P}{\partial t} \;=\; -\cos\theta \frac{\partial P}{\partial x} \;-\; \sin\theta \frac{\partial P}{\partial y} \;+\; \frac{\sigma^2}{2} \frac{\partial^2 P}{\partial \theta^2} \;-\; \frac{1}{\tau} P \; , \tag{4.3}$$

where the initial probability distribution of sources (or sinks) is described by $P(\boldsymbol{x}, \theta\,; 0)$. This partial differential equation can be viewed as a set of independent *advection* equations in $\boldsymbol{x} = (x, y)$ (the first and second terms) coupled in the θ dimension by the *diffusion* equation (the third term). The advection equations translate probability mass in direction θ with unit speed, while the diffusion term models the Brownian motion in direction, with *diffusion parameter*, σ. The combined effect of these three terms 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. Finally, the effect of the fourth term is that particles decay over time, with a half life given by the *decay constant*, τ. This represents our prior expectation on the length of gaps—most are quite short. In [22] stochastic completion fields were computed by solving the Fokker-Planck equation using a standard finite differencing scheme on a regular grid.

5 Description of Algorithm

One of the main goals of this paper is to derive a discrete numerical algorithm to compute stochastic completion fields in a shift-twist invariant manner. This invariance is achieved by first evolving the Fokker-Planck equation in a shiftable-twistable basis of $\mathbf{R}^2 \times S^1$ to obtain representations of the source and sink fields in the basis, and then multiplying these representations in a shift-twist invariant manner to obtain a representation of the completion field in a shiftable-twistable basis.

We observe that a discrete Dirac basis, consisting of functions, $\Psi_{\boldsymbol{k},m}(\boldsymbol{x},\theta) = \delta(\boldsymbol{x} - \boldsymbol{k}\Delta)\,\delta(\theta - m\Delta_\theta)$, where $(\boldsymbol{k}, m)$ is a triple of integers, is not shiftable-twistable. This is because a Dirac function located off the grid of Dirac basis functions is not in their span.

A major shortcoming of previous contour completion algorithms [8, 9, 13, 16, 17, 21, 22, 25] is that they perform computations in this basis. As a consequence, initial conditions which do not lie directly on the grid cannot be accurately represented. This problem is often skirted by researchers in this area by choosing input patterns which match their choice of sampling rate and phase. For example, Li [13] used only six orientations (including $0°$) and Heitger and von der Heydt [9], only twelve (including $0°$, $60°$ and $120°$). Li's first test pattern was a line of orientation, $0°$, while Heitger and von der Heydt used a Kanizsa Triangle with sides of $0°$, $60°$, and $120°$ orientation. There is no reason to believe that the experimental results they show would be the same if their input patterns were rotated by as little as $5°$.[5]

In addition to the problem of representing the input, the computation itself must be Euclidean invariant. Stochastic completion fields computed using the finite differencing scheme of [22] exhibit marked anisotropic spatial smoothing due to the manner in which 2D advection is performed on a grid (see Figures 4,5 and 6). Although probability mass advects perfectly in either of the two principal coordinate directions, mass which is moving at an angle to the grid gradually disperses, since, at each time step, bilinear interpolation is used to place the mass on the grid.

For reasons of simplicity, in this paper, we chose to compute stochastic completion fields in a Gaussian-Fourier basis.[6] The initial conditions for the Fokker-

[5] Nor are we blameless in this respect. Williams and Jacobs [21, 22] used 36 directions (including $0°$, $60°$ and $120°$) and demonstrated their computation with a Kanizsa Triangle with sides of $0°$, $60°$ and $120°$ orientation.

[6] The computation can also be performed in more biologically plausible shiftable-twistable bases, the simplest of which is the CDDG-Fourier basis.

Planck initial value problem are modeled by fine scale, three-dimensional Gaussians, whose centers are determined by the locations and directions of the edge fragments to be completed. We use the single scale method discussed in Section 3 to represent the initial conditions in the basis.

To solve the Fokker-Planck equation, we express its solution in terms of the basis functions, $G_{\boldsymbol{k},\omega}(\boldsymbol{x},\theta)$, as

$$P(\boldsymbol{x},\theta\,;t) \;=\; \textstyle\sum_{\boldsymbol{k},\omega}\; c_{\boldsymbol{k},\omega}(t)\, G_{\boldsymbol{k},\omega}(\boldsymbol{x},\theta)\ , \tag{5.1}$$

where the coefficients, $c_{\boldsymbol{k},\omega}(t)$, depend on time. Then, we derive a linear transformation, $c(t+\Delta t) = (\mathbf{A}\circ\mathbf{D})c(t)$, to evolve the coefficient vector in time. This transformation is the composition of an advection transformation, $\mathbf{A}$, which has the effect of transporting probability mass in directions θ, and a diffusion-decay transformation, $\mathbf{D}$, which implements both the diffusion of mass in θ, and the decay of mass over time. Representations of source or sink fields in the basis are obtained by integrating the coefficient vector, $c(t)$, over time, where the initial coefficient vector represents the initial sources or sinks.

The shiftability-twistability of the basis functions is used in two distinct ways to obtain shift-twist invariant source and sink fields. First, it enables any two initial conditions, which are related by an arbitrary transformation, $T_{\boldsymbol{x}_0,\theta_0}$, to be represented equally well in the basis. Second, it is used to derive a shift-twist invariant advection transformation, $\mathbf{A}$, thereby eliminating the grid orientation artifacts described above. In summary, given a desired resolution at which to represent the initial conditions, our new algorithm produces source and sink fields, at the given resolution, which transform appropriately under arbitrary Euclidean transformations of the input image. In contrast, in all previous contour completion algorithms, the degree of failure of Euclidean invariance is highly dependent on the resolution of the grid, and can be quite large relative to the grid resolution.

The final step in our shift-twist invariant algorithm is to compute the completion field (the product of the source and sink fields) in a shiftable-twistable basis. The particular basis used to represent completion fields is the same as the one used to represent the source and sink fields, except that the variance of the Gaussian basis functions in $\mathbf{R}^2$ needs to be halved. The need to use a slightly different basis to represent completion fields is not biologically implausible, since the experimental evidence described in Section 2 suggests that the neural locus of the source and sink fields could be V1, while completion fields are more likely located in V2.

6 The Solution of the Fokker-Planck Equation

In this section we derive a shift-twist invariant linear transformation, $c(t+\Delta t) = (\mathbf{A}\circ\mathbf{D})c(t)$, of the coefficient vector which evolves the Fokker-Planck equation in a shiftable-twistable basis. The derivation holds for any shiftable-twistable basis constructed from shiftable-twistable functions of the form, $\Psi_\omega(\boldsymbol{x},\theta) = \psi(\boldsymbol{x})e^{i\omega\theta}$, for some function, $\psi(\boldsymbol{x})$. Since the transformation, $\mathbf{A} \circ \mathbf{D}$, will only involve interactions between functions, $\psi(\boldsymbol{x})$, at different positions $\boldsymbol{k}\Delta$, and not at different scales or orientations, the basis functions and coefficients will be denoted by $\Psi_{\boldsymbol{k},\omega}$ and $c_{\boldsymbol{k},\omega}(t)$ respectively.[7]

To derive an expression,

$$c_{\boldsymbol{\ell},\eta}(t + \Delta t) \;=\; \textstyle\sum_{\boldsymbol{k},\omega} \; \mathbf{A}_{\boldsymbol{\ell},\eta\,;\,\boldsymbol{k},\omega}(\Delta t)\; c_{\boldsymbol{k},\omega}(t) \;, \tag{6.1}$$

for the advection transformation, $\mathbf{A}$, in the basis, $\Psi_{\boldsymbol{k},\omega}$, we exploit the fact that spatial advection can be done perfectly using shiftable basis functions, $\psi_{\boldsymbol{k}}(\boldsymbol{x})$, in $\mathbf{R}^2$, and the continuous variable, $\theta \in S^1$. Suppose that P is given in the form,

$$P(\boldsymbol{x},\theta;\, t) \;=\; \textstyle\sum_{\boldsymbol{k},\omega} \; c_{\boldsymbol{k},\omega}(t)\; \psi_{\boldsymbol{k}}(\boldsymbol{x})\, e^{i\omega\theta} \;=\; \sum_{\boldsymbol{k}} \; \check{c}_{\boldsymbol{k},\theta}(t)\; \psi_{\boldsymbol{k}}(\boldsymbol{x}) \;, \tag{6.2}$$

where $\check{c}(t)$ is related to $c(t)$ by the standard synthesis formula for Fourier series, $\check{c}_{\boldsymbol{k},\theta} = \sum_\omega c_{\boldsymbol{k},\omega} e^{i\omega\theta}$, which we denote by $\check{c} = \mathbf{F}^{-1}c$. Then the translation of P in direction, θ, at unit speed, for time, Δt, is given by

$$P(\boldsymbol{x},\theta\,; t + \Delta t) \;=\; P(\boldsymbol{x} - \Delta t[\cos\theta, \sin\theta]^{\mathrm{T}}, \theta\,; t) \tag{6.3}$$

$$\;=\; \sum_{\boldsymbol{k}} \; \check{c}_{\boldsymbol{k},\theta}(t)\; \psi_{\boldsymbol{k}}(\boldsymbol{x} - \Delta t[\cos\theta, \sin\theta]^{\mathrm{T}}) \;, \tag{6.4}$$

where the second equation follows from equation (6.2). The shiftability of ψ then implies that

$$\check{c}_{\boldsymbol{\ell},\theta}(t + \Delta t) \;=\; \textstyle\sum_{\boldsymbol{k}} \; \check{\mathbf{A}}_{\boldsymbol{\ell},\theta;\boldsymbol{k},\theta}(\Delta t)\; \check{c}_{\boldsymbol{k},\theta}(t) \;, \tag{6.5}$$

where

$$\check{\mathbf{A}}_{\boldsymbol{\ell},\theta;\boldsymbol{k},\theta}(\Delta t) = b_{\boldsymbol{\ell}-\boldsymbol{k}}(\Delta t[\cos\theta, \sin\theta]^{\mathrm{T}}) \;. \tag{6.6}$$

Finally, the advection transformation, $\mathbf{A}$, in the basis, $\Psi_{\boldsymbol{k},\omega}$, is given by the similarity transformation, $\mathbf{A} = \mathbf{F}\check{\mathbf{A}}\mathbf{F}^{-1}$, where $\mathbf{F}$ denotes the standard analysis formula for Fourier series, $(\mathbf{F}f)(\omega) = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)e^{-i\omega\theta}d\theta$. Since $c = \mathbf{F}\check{c}$ we have the following result.

Theorem 1. *In the basis, $\Psi_{\boldsymbol{k},\omega}$, the advection transformation, $\mathbf{A}$, is given by*

$$c_{\boldsymbol{\ell},\eta}(t + \Delta t) \;=\; \textstyle\sum_{\boldsymbol{k},\omega} \; \hat{b}_{\boldsymbol{\ell}-\boldsymbol{k},\eta-\omega}(\Delta t)\; c_{\boldsymbol{k},\omega}(t) \;, \tag{6.7}$$

where

$$\hat{b}_{\boldsymbol{k},\eta}(\Delta t) \;=\; \tfrac{1}{2\pi}\textstyle\int_0^{2\pi} b_{\boldsymbol{k}}(\Delta t[\cos\theta, \sin\theta]^{\mathrm{T}})\, e^{-i\eta\theta}\, d\theta \;. \tag{6.8}$$

In particular, the transformation, $\mathbf{A}$, is shift-twist invariant and is a convolution operator on the vector space of coefficients, $c_{\boldsymbol{k},\omega}$.

[7] Since we are using Fourier series in θ the transformation, $\mathbf{D}$, can be implemented in a shift-twist invariant manner by applying a standard finite differencing scheme to the coefficients.

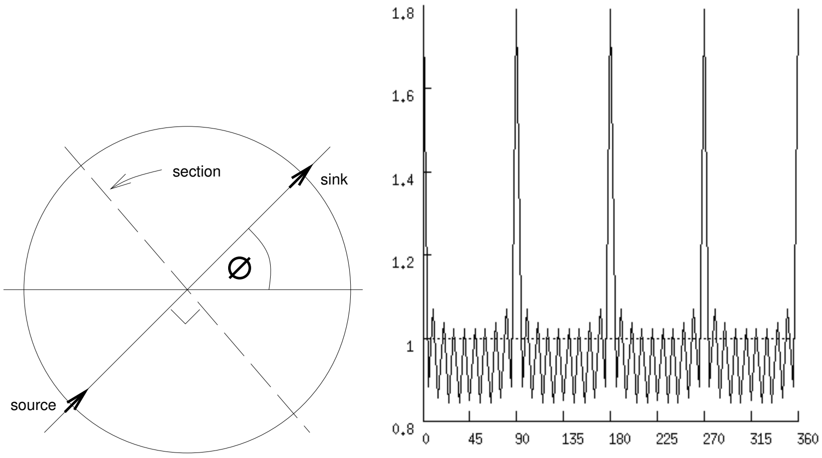

Fig. 3. The geometry of the straight line completion field experiment (*left*). Graph (*right*) of the mean along a section normal to the straight line completion field as a function of the direction, ϕ, for our new algorithm (*dashed line*) and for the algorithm of [22] (*solid line*)

Theorem 1 implies that the computation of source and sink fields can be performed in a recurrent neural network using a fixed set of units as described in [22]. Since the advection transformation, $\mathbf{A}$, is a convolution operator on the space of coefficients, for efficiency's sake we implemented both $\mathbf{A}$ and $\mathbf{D}$ in the 3D Fourier domain of the coefficient vector. In this domain, $\mathbf{A}$ is given by a diagonal matrix and $\mathbf{D}$ by a circulant tridiagonal matrix.

7 Experimental Results

We present three experiments demonstrating the Euclidean invariance of our algorithm. In each experiment, the Gaussian-Fourier basis consisted of $K = 160$ translates in each spatial variable of a Gaussian (of period $X = 40.0$ units), and harmonic signals of $N = 92$ frequencies in the angular variable, for a total of 2.355×10^6 basis functions. Pictures of completion fields were obtained by analytically integrating over θ and rendering the completion field on a 256×256 grid.

We compare the new algorithm with the finite differencing scheme of [22]. For the method of [22], the $40.0 \times 40.0 \times 2\pi$ space was discretized using a 256×256 spatial grid with 36 discrete orientations, for a total of 2.359×10^6 Dirac basis functions. The intent was to use approximately the same number of basis functions for both algorithms. The initial conditions were represented on the grid using tri-linear interpolation and pictures of the completion fields were obtained by summing over the discrete angles. The same parameters were used for both algorithms. The decay constant was $\tau = 4.5$ and the time increment, $\Delta t = 0.1$.

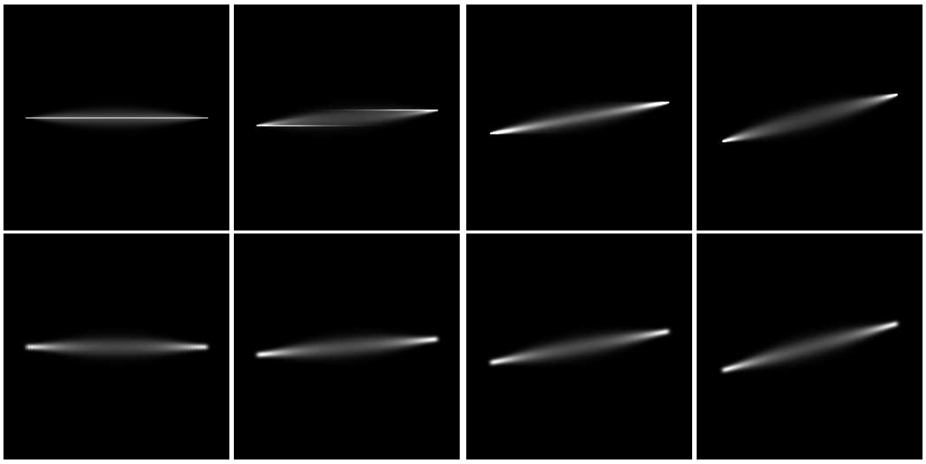

Fig. 4. Straight line completion fields due to an initial stimulus consisting of two points on a circle with direction, ϕ, normal to the circle, for $\phi = 0°, 5°, 10°, 15°$ (*left to right*). The completion fields were computed using the algorithm of [22] (*top row*) and using the new algorithm (*bottom row*)

The diffusion parameter was $\sigma = 0.08$ for the first and second experiments and $\sigma = 0.14$ for the third.[8] In Figures 4, 5 and 6 the completion fields constructed using the algorithm of [22] are in the top row, while those constructed using the new algorithm are in the bottom row.

In the first experiment, we computed straight line completion fields joining two diametrically opposed points on a circle of radius, 16.0, with initial directions normal to the circle. That is, given an angle, ϕ, the initial stimulus consisted of the two points, $(\pm 16.0 \cos\phi, \pm 16.0 \sin\phi, \phi)$, see Figure 3 (left). The completion fields are shown in Figure 4, with those in the top row, computed using the method of [22], clipped above at 2×10^{-6}.

To compare the degree of Euclidean invariance of the two algorithms, we extracted a section of each completion field along the diameter of the circle normal to the direction of the completion field. In Figure 3 (right), we plot the mean of each section as a function of the angle ϕ. The dashed line indicates the means computed using the new algorithm, and the solid line shows the means computed using the algorithm of [22].[9] The fact that the dashed line graph is constant provides solid evidence for the Euclidean invariance of the new algorithm. The solid line graph demonstrates the two major sources of the lack of Euclidean invariance in the method of [22]. First, the rapid oscillation of

[8] The diffusion parameter, σ, was required to be larger in the third experiment because of the high curvature circles in the Kanizsa triangle figure.

[9] The angles, ϕ, were taken in $5°$ increments from $0°$ to $45°$. For illustration purposes the ϕ-axis was extended to $360°$ so as to reflect the symmetry of the grid. Both graphs were normalized to have average value one.

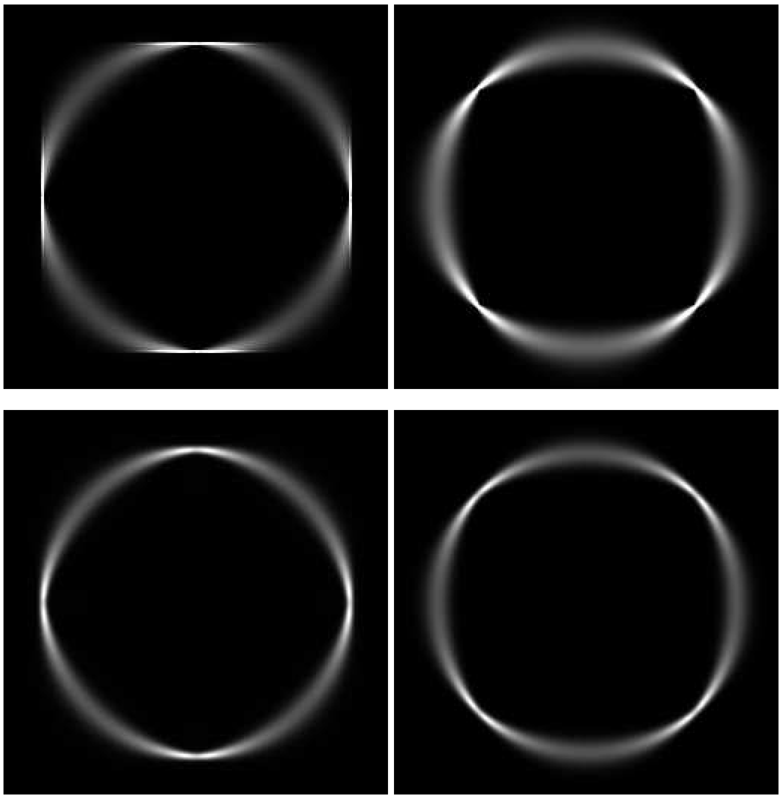

Fig. 5. Completion fields due to the Ehrenstein initial stimulus in Figure 1(a) (*left column*) and with the initial conditions rotated clockwise by 45° (*right column*). The completion fields were computed using the algorithm of [22] (*top*) and using the new algorithm (*bottom*)

period 10° is due to the initial conditions coming in and out of phase with the angular grid. This 10° periodicity can be seen in the periodicity of the general shape of the completion fields in the top row of Figure 4. Second, the large spikes at 90° intervals are due to the anisotropic manner in which the advection transformation was solved on the spatial grid. These large spikes correspond to the very bright horizontal line artifacts in the first two completion fields in the top row of Figure 4.

In the second experiment, we computed completion fields due to rotations of the Ehrenstein initial stimulus in Figure 1(a). Pictures of the completion fields are shown in Figure 5.[10] The left column shows the completion fields due to the Ehrenstein stimulus in Figure 1(a), while in the right column the

[10] Because of the periodicity in the spatial variables, $\boldsymbol{x}$, to avoid wrap around in this experiment, for the new algorithm the computation was performed on a $80.0 \times 80.0 \times 2\pi$ space with $K = 320$.

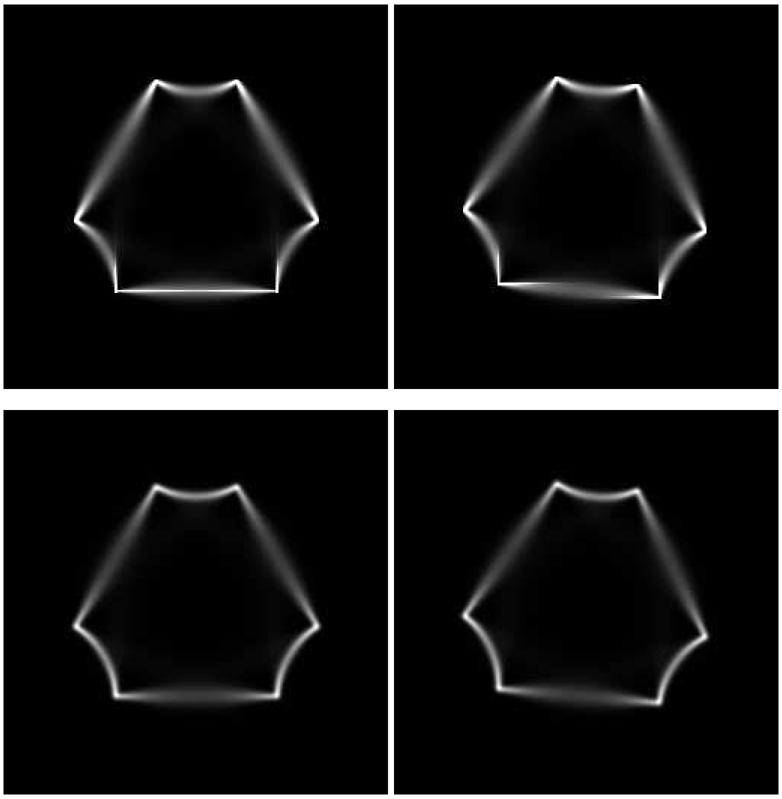

Fig. 6. Completion fields due to the Kanizsa triangle initial stimulus in Figure 1(b) (*left column*) and with the initial conditions rotated clockwise by $5°$ (*right column*). The completion fields were computed using the algorithm of [22] (*top*) and using the new algorithm (*bottom*)

initial conditions have been rotated clockwise by $45°$. The completion fields computed using the method of [22] were clipped above at 1.25×10^{-8}. For our final experiment, we compute completion fields due to rotations and translations of the Kanizsa Triangle stimulus in Figure 1(b). Completion fields are shown in Figure 2, which was discussed in the Introduction, and in Figure 6. The left column of Figure 6 shows completion fields due to the Kanizsa Triangle in Figure 1(b). In the right column the initial conditions have been rotated clockwise by $5°$. The completion fields computed using the method of [22] were clipped above at 9×10^{-5}.

The completion fields in the bottom rows of Figures 5 and 6, and in Figure 2, demonstrate the Euclidean invariance of our new algorithm. This is in marked contrast with the obvious lack of Euclidean invariance in the completion fields in the top rows of Figures 5 and 6. The visible straight line artifacts in these completion fields, which are oriented along the coordinate axes, are due to the anisotropic nature of the advection process in the algorithm of [22], and (to a lesser extent), to the way in which the initial conditions were represented on the grid.

8 Conclusion

An important initial stage in the analysis of a scene requires completion of the boundaries of partially occluded objects. Williams and Jacobs introduced the notion of the stochastic completion field which measures the probability distribution of completed boundary shapes in a given scene. In this article we have described a new, parallel, algorithm for computing stochastic completion fields. As is required of any computational model of human visual information processing, our algorithm attempts to reconcile the apparent contradiction between the Euclidean invariance of human early visual computations on the one hand, and the observed sparseness of the discrete spatial sampling of the visual field by primary and secondary visual cortex on the other hand. The new algorithm reconciles these two contradictions by performing the computation in a basis of separable functions with spatial components similar to the receptive fields of simple cells in primary visual cortex. In particular, the Euclidean invariance of the computation is achieved by exploiting the shiftability and twistability of the basis functions.

In this paper, we have described three basic results. First, we have generalized Simoncelli *et al.'s* notion of shiftability and steerability in $\mathbf{R}^2$ to a more general notion of shiftability and twistability in $\mathbf{R}^2 \times S^1$. The notion of shiftability and twistability mirrors the coupling between the advection and diffusion terms in the Fokker-Planck equation, and at a deeper level, basic symmetries in the underlying random process characterizing the distribution of completion shapes. Second, we described a new method for numerical solution of the Fokker-Planck equation in a shiftable-twistable basis. Finally, we used this solution to compute stochastic completion fields, and demonstrated, both theoretically and experimentally, the invariance of our computation under translations and rotations of the input pattern.

References

1. Blasdel, G., and Obermeyer, K., Putative Strategies of Scene Segmentation in Monkey Visual Cortex, *Neural Networks*, **7**, pp. 865-881, 1994.
2. Cowan, J.D., Neurodynamics and Brain Mechanisms, *Cognition, Computation and Consciousness*, Ito, M., Miyashita, Y. and Rolls, E., (Eds.), Oxford UP, 1997.
3. Daugman, J., Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-dimensional Visual Cortical Filter, *J. Opt. Soc. Am. A*, **2**, pp.1160-1169, 1985.
4. Daugman, J., Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression, *IEEE Trans. Acoustics, Speech, and Signal Processing* **36**(7), pp. 1,169-1,179, 1988.

5. Eyesel, U. Turning a Corner in Vision Research, *Nature*, **399**, pp. 641-644, 1999.
6. Freeman, W., and Adelson, E., The Design and Use of Steerable Filters, IEEE Trans. PAMI, **13** (9), pp.891-906, 1991.
7. Gilbert, C.D., Adult Cortical Dynamics, *Physiological Review*, **78**, pp. 467-485, 1998.
8. Grossberg, S., and Mingolla, E., Neural Dynamics of Form Perception: Boundary Completion, Illusory Figures, and Neon Color Spreading, *Psychological Review*, **92**, pp. 173-211, 1985.
9. Heitger, R. and von der Heydt, R., A Computational Model of Neural Contour Processing, Figure-ground and Illusory Contours, *Proc. of 4th Intl. Conf. on Computer Vision*, Berlin, Germany, 1993.
10. von der Heydt, R., Peterhans, E. and Baumgartner, G., Illusory Contours and Cortical Neuron Responses, *Science*, **224**, pp. 1260-1262, 1984.
11. Iverson, L., Toward Discrete Geometric Models for Early Vision, Ph.D. dissertation, McGill University, 1993.
12. Kalitzin, S., ter Haar Romeny, B., and Viergever, M., Invertible Orientation Bundles on 2D Scalar Images, in *Scale-Space Theory in Computer Vision*, ter Haar Romeny, B., Florack, L., Koenderink, J. and Viergever, M., (Eds.), Lecture Notes in Computer Science, **1252**, 1997, pp. 77-88.
13. Li, Z., A Neural Model of Contour Integration in Primary Visual Cortex, *Neural Computation*, **10**(4), pp. 903-940, 1998.
14. Marčelja, S. Mathematical Description of the Responses of Simple Cortical Cells, *J. Opt. Soc. Am.*, **70**, pp. 1297-1300, 1980.
15. Mumford, D., Elastica and Computer Vision, *Algebraic Geometry and Its Applications*, Chandrajit Bajaj (ed.), Springer-Verlag, New York, 1994.
16. Parent, P., and Zucker, S.W., Trace Inference, Curvature Consistency and Curve Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, pp. 823-889, 1989.
17. Shashua, A. and Ullman, S., Structural Saliency: The Detection of Globally Salient Structures Using a Locally Connected Network, *2nd Intl. Conf. on Computer Vision*, Clearwater, FL, pp. 321-327, 1988.
18. Simoncelli, E., Freeman, W., Adelson E. and Heeger, D., Shiftable Multiscale Transforms, *IEEE Trans. Information Theory*, **38**(2), pp. 587-607, 1992.
19. Thornber, K.K. and Williams, L.R., Analytic Solution of Stochastic Completion Fields, *Biological Cybernetics* **75**, pp. 141-151, 1996.
20. Thornber, K.K. and Williams, L.R., Orientation, Scale and Discontinuity as Emergent Properties of Illusory Contour Shape, *Neural Information Processing Systems* **11**, Denver, CO, 1998.
21. Williams, L.R., and Jacobs, D.W., Stochastic Completion Fields: A Neural Model of Illusory Contour Shape and Salience, *Neural Computation*, **9**(4), pp. 837-858, 1997, (also appeared in *Proc. of the 5th Intl. Conference on Computer Vision (ICCV '95)*, Cambridge, MA).
22. Williams, L.R., and Jacobs, D.W., Local Parallel Computation of Stochastic Completion Fields, *Neural Computation*, **9**(4), pp. 859-881, 1997.
23. Williams, L.R. and Thornber, K.K., A Comparison of Measures for Detecting Natural Shapes in Cluttered Backgrounds, *Intl. Journal of Computer Vision*, **34** (2/3), pp. 81-96, 1999.
24. Wandell, B.A., *Foundations of Vision*, Sinauer Press, 1995.
25. Yen, S. and Finkel, L., Salient Contour Extraction by Temporal Binding in a Cortically-Based Network, *Neural Information Processing Systems* **9**, Denver, CO, 1996.