Boston University

OpenBU **http://open.bu.edu**

Cognitive & Neural Systems CAS/CNS Technical Reports

1996-08

Real-Time Anisotropic Diffusion using Space-Variant Vision

https://hdl.handle.net/2144/2325

"Downloaded from OpenBU. Boston University's institutional repository."

REAL-TIME ANISOTROPIC DIFFUSION USING SPACE-VARIANT VISION

Bruce Fischl, Eric L. Schwartz, and Michael A. Cohen

August 1996

Technical Report CAS/CNS-96-026

Permission to copy without fee all or part of this material is granted provided that: 1. the copies are not made or distributed for direct commercial advantage, 2. the report title, author, document number, and release date appear, and notice is given that copying is by permission of the BOSTON UNIVERSITY CENTER FOR ADAPTIVE SYSTEMS AND DEPARTMENT OF COGNITIVE AND NEURAL SYSTEMS. To copy otherwise, or to republish, requires a fee and/or special permission.

Copyright © 1996

Boston University Center for Adaptive Systems and
Department of Cognitive and Neural Systems
677 Beacon Street
Boston, MA 02215

Real-Time Anisotropic Diffusion using Space-Variant Vision

Bruce Fischl†

Dept. Cognitive and Neural Systems
Boston University
677 Beacon Street
Boston, MA 02215
email: fischl@cns.bu.edu

Eric L. Schwartz†*

Dept. Cognitive and Neural Systems
Boston University
677 Beacon Street
Boston, MA 02215
email: eric@thing4.bu.edu
(617) 353-6179

Michael A. Cohen

Dept. Cognitive and Neural Systems
Boston University
677 Beacon Street
Boston, MA 02215
email: mike@cns.bu.edu

* Author to whom reprint requests should be made.
† Supported in part by the office of naval research (ONR N00014-95-1-0409).

Abstract - *Many computer and robot vision applications require multi-scale image analysis. Classically, this has been accomplished through the use of a linear scale-space, which is constructed by convolution of visual input with Gaussian kernels of varying size (scale). This has been shown to be equivalent to the solution of a linear diffusion equation on an infinite domain, as the Gaussian is the Green's function of such a system (Koenderink, 1984). Recently, much work has been focused on the use of a variable conductance function resulting in anisotropic diffusion described by a nonlinear partial differential equation (PDE). The use of anisotropic diffusion with a conductance coefficient which is a decreasing function of the gradient magnitude has been shown to enhance edges, while decreasing some types of noise (Perona and Malik, 1987). Unfortunately, the solution of the anisotropic diffusion equation requires the numerical integration of a nonlinear PDE which is a costly process when carried out on a fixed mesh such as a typical image. In this paper we show that the complex log transformation, variants of which are universally used in mammalian retino-cortical systems, allows the nonlinear diffusion equation to be integrated at exponentially enhanced rates due to the non-uniform mesh spacing inherent in the log domain. The enhanced integration rates, coupled with the intrinsic compression of the complex log transformation, yields a speed increase of between two and three orders of magnitude, providing a means of performing real-time image enhancement using anisotropic diffusion.*

Keywords: anisotropic diffusion, space-variant vision, log-polar, image enhancement.

1. Introduction.

Multi-scale image enhancement and representation is an important part of biological and machine early vision systems. The process of constructing this representation must be both rapid and insensitive to noise, while retaining image structure at all scales. This is a complex problem as small-scale structure is difficult to distinguish from noise, while larger scale structure requires more computational effort, and can be hard to accurately localize. Errors can also arise when conflicting results at different scales require cross-scale arbitration.

Attempts to solve problems of this type resulted in the linear scale-space formulation of Witken (Witken, 1983) in which an image is convolved with Gaussian kernels of various sizes. Edges delineating the boundaries between objects can then be found in a number of ways, for example by tracing the zero-crossings of the Laplacian through scale-space, similar to the manner proposed by Marr and Hildreth (Marr and Hildreth, 1980). This approach is problematic as the zeros can change position and disappear as scale-space is traversed due to the influence of neighboring structure. In this situation it is unclear how to arbitrate between conflicting results at different scales.

Koenderink (Koenderink, 1984) and Hummel (Hummel, 1986) pointed out that the one-parameter family of images comprising scale-space can be equivalently viewed as snapshots of the time-evolution of the diffusion (or heat) equation:

$$I_t = c\Delta I \tag{1.1}$$

Where I is the intensity image, c is a diffusion constant, I_t is the partial derivative of I with respect to time, and Δ is the Laplacian operator with respect to the spatial coordinates.

The diffusion equation provides a mathematical framework with which to analyze the scale-space formalism, but it does not address the issue of cross-scale comparison. While Koenderink restricted his analysis to isotropic diffusion characterized by the linear heat equation, Perona and Malik (Perona and Malik, 1987, Perona and Malik, 1990) suggested that a nonlinear anisotropic version of the heat equation could remedy some of the difficulties encountered in the use of a linear scale-space. This followed earlier psychophysical and neurophysiological modelling work which used variable diffusion to account for a variety of human perceptual phenomena (Cohen and Grossberg, 1984; Grossberg and Mingolla, 1985). Perona and Malik proposed the following equation in which the conduction coefficient is not constant in space, but is rather a function of the magnitude of the intensity gradient of the image:

$$I_t = \nabla \cdot (c(|\nabla I|)\nabla I) \tag{1.2}$$

In this way, the amount of diffusion at each point in space is modulated by the function $c(|\nabla I|)$, and the image gradient at that point. They choose to make $c(\cdot)$ a decreasing function of the image gradient magnitude, so that regions of high contrast undergo less diffusion, and are therefore preserved over time. This is in contrast to the linear heat equation which blurs uniformly, destroying small scale structure as time evolves. Systems such as equation (1.2) are intended to yield a single intensity image which retains edge information at all scales of interest, thus obviating the need for any type of cross-scale arbitration.

The Perona-Malik equation (1.2) is a nonlinear partial differential equation of a type which is difficult to analyze. It has been suggested (Nitzberg and Shiota, 1992) that (1.2) is unstable for some parameter regimes, although this is still a point of investigation (Perona et al., 1994). Furthermore, it can amplify small scale noise which gives rise to high gradient magnitudes. Many variants of the Perona and Malik scheme have been proposed to improve its sensitivity to noise, its speed, its instability, and its equilibrium behavior (Alvarez et al., 1992; Alvarez and Mazorra, 1994; Catte et al., 1992; Cottet and Germain, 1993; Dang et al., 1994; El-Fallah and Ford, 1994; Engquist et al., 1989; Illner and Neunzert, 1993; Li and Chen, 1994; Nitzberg and Shiota, 1992; Nördstrom, 1990; Osher and Rudin, 1990; Pauwels et al., 1993; Price et al., 1990; Whitaker and Pizer, 1991;Whitaker, 1993; Kacur and Mikula, 1995; Fischl and Schwartz, 1995; Fischl and Schwartz, 1996; Shah, 1996; Malladi and Sethian, 1995).

The diffusion paradigm, while impressive in the quality of the images it produces, suffers from a number of drawbacks. The most prominent of these is the computational cost of the algorithms, coupled with their inherently serial nature. This makes them implausible from a biological standpoint, as well as impractical for use in real-time machine vision applications. The biological implausibility stems from the relatively rapid nature of perception relative to neural conduction delays and peak firing rates (≤200Hz). Psychophysical and neurophysiological experiments indicate that perception can occur as rapidly as 100-150 msec (Thorpe and Imbert, 1989, Oram and Perrett, 1992) which is only 40 milliseconds or so longer than the latency of cells in primary visual cortex (Vogels and Orban, 1991). Using these figures together with typical firing frequencies and synaptic transmission delays, Thorpe and Imbert (Thorpe and Imbert, 1989) argue that the number of synaptic connections (assumed to be equivalent to the number of serial steps) used by the visual system in rapid identification tasks is somewhere between 10 and 50, although probably closer to the lower bound. Thus, while complex processing is possible, it is almost certainly parallel in nature, ruling out numerical schemes which require more than a few iterations.

Almost without exception the use of anisotropic diffusion in machine and biological vision research has been performed in the space-invariant or Cartesian domain. However, it has been shown that the mapping from the mammalian retina to striate cortex is a space-variant one which can be well approximated by a complex log transformation (Schwartz, 1977; Schwartz, 1980; Schwartz, 1994). Despite some notable advantages (dramatic pixel count reduction, quasi size and rotation invariance), the complex log map has not been widely used in the machine vision community. In large part this has been due to the lack of shape invariance under translation, which severely complicates object recognition. This drawback has recently been addressed (Bonmassar and Schwartz, 1995), allowing frequency domain techniques to be applied in the complex log domain.

In this paper we derive the form that the nonlinear diffusion equation takes in the space-variant coordinate system, and show that it has a number of notable advantages. Most importantly, the nonuniform pixel spacing inherent in the log domain allows integration to proceed at a nonuniform rate which is an exponential function of the radial coordinate. Thus, the peripheral parts of a log plane image move rapidly through scale-space, achieving large-scale image enhancement in dramatically fewer time steps than the corresponding process in Cartesian space. The reduction in integration time coupled with the compressive effect of the complex log transformation itself yields more than two orders of magnitude speed increase for log plane diffusion, allowing image enhancement to take place in real-time.

2. Space-Variant Vision.

The mammalian retina is a space-variant sensor: the spacing of sensory neurons across the retinal surface is not uniform. The density of cells is greatest in the high acuity fovea, and falls off with retinal eccentricity. This allows the simultaneous achievement of high resolution and a wide field of view without requiring an enormous number of sensing elements. This anatomical feature has clear perceptual correlates. Visual acuity in the fovea is greater than in the periphery by at least a factor of 40 (Wertheim, 1894). This is the result of many factors including the optics of the eye (Campbell and Green, 1965), photoreceptor sampling density (Williams and Coletta, 1987), spatial averaging due to the size of peripheral receptive fields (Merigan and Katz, 1990), as well as ganglion cell density (Wässle et al., 1990).

The mapping from the retina to striate cortex has been shown to be well approximated by a complex log map (Schwartz, 1977; Schwartz, 1980). This discovery has motivated the use of the complex log mapping in the construction of space variant sensors and algorithms for machine vision systems (Rojer and Schwartz, 1990; Weiman, 1988; Sandini and Dario, 1989, Sandini et al., 1989; Messner and Szu, 1986; Schenker et al., 1981; Bonmassar and Schwartz, 1994; Bonmassar and Schwartz, 1995; Bonmassar and Schwartz, 1996b; Bonmassar and Schwartz, 1996a; Yamamoto et al., 1996). The log mapping expresses the variation in cortical area devoted to different regions of the retina.

2.1. Space-Variant Vision in Biology.

The investigation of the space-variant properties of the mammalian retino-cortical mapping dates back to the early 1960's. In order to characterize the transformation of visual data from retinal coordinates to primary visual cortex Daniel and Whitteridge introduced the concept of the cortical magnification factor M_c, measured in millimeters of cortex per degree of visual angle (Daniel and Whitteridge, 1961). The magnification factor is not constant across the retina, but rather varies as a function of eccentricity. Experimentally, the cortical magnification factor has been found to be accurately approximated by (Wilson et al., 1990)

$$M_c(r) = \frac{C_1}{1 + C_2 r} \tag{2.1}$$

Where r is the retinal eccentricity measured in degrees, and C_1 and C_2 are experimentally determined constants related to the foveal magnification and the rate at which magnification falls off with eccentricity. Integrating equation (2.1) yields a relationship between retinal eccentricity and cortical distance ρ

$$\rho(r) = \int_0^r \frac{C_1}{1 + C_2 r'} dr' = \frac{C_1}{C_2} \log(1 + C_2 r) \tag{2.2}$$

Schwartz (Schwartz, 1977; Schwartz, 1980) has pointed out that the cortical magnification factor should be considered a vector quantity as opposed to a scalar one. The retino-cortical mapping can then be conveniently and concisely expressed as a conformal transformation. In this approach, a complex variable z is used to describe the retinal coordinates

$$z = re^{i\theta} = x + iy \tag{2.3}$$

Where polar coordinates replace Cartesian ones in the retina

$$r = \sqrt{x^2 + y^2}, \ \theta = \tan^{-1}\left(\frac{y}{x}\right) \tag{2.4}$$

The cortical point $(\rho(z), \phi(z))$ can then be specified by a single complex variable w as

$$w = \rho(z) + i\phi(z) = K\log(z + a), \quad Re(z) \geq 0 \tag{2.5}$$

Where K is a scale factor determined by cortical area, which will be dropped in the following discussion, and a is a real positive constant, called the map parameter. The value of a determines the size of the quasi-linear region around z=0, and is generally believed to be in the range 0.3 to 0.7 degrees (see (Schwartz, 1994) for a discussion of the significance of K and a). The effect of modifying a on the mapping can be seen in the following way (see figure (2.1)). For small z (i.e. $z << a$), the mapping can be approximated using a series expansion around the point z=0:

$$w \approx \log(a) + \frac{z}{a} \tag{2.6}$$

Thus, in the fovea, the mapping is essentially linear. The magnitude of the derivative of the mapping gives an approximation to the cortical magnification factor:

$$\left|\frac{dw}{dz}\right| = \left|\frac{1}{z + a}\right| \tag{2.7}$$

Which is approximately constant for $z << a$. The complex log transformation of equation (2.5) therefore smoothly varies from a linear map in the fovea to a logarithmic one in the periphery, with the magnitude of a controlling the size of the region of approximate linearity. This is in contrast with other techniques which explicitly overlay a Cartesian fovea on a log image to obtain a similar effect (Sandini and Dario, 1989; Sandini et al., 1989).

Equation (2.5) is analytic everywhere in the domain and is hence conformal, implying that local angles are preserved by the transformation (Churchill and Brown, 1984). The singularity at the origin for the more commonly used complex log mapping $w=log(z)$ is removed at the cost of mapping the two hemifields separately and managing a discontinuity along the vertical meridian. The full form of the mapping for both hemifields[1] is given by

$$w = \begin{cases} \log(z+a) & , \quad \mathrm{Real}(z) \geq 0 \\ 2\log(a) - \log(-z+a) & , \quad \mathrm{Real}(z) < 0 \end{cases} \tag{2.8}$$

Figure (2.1) shows an example of an image, and its complex log transformation for a variety of values of the map parameter a. As can be seen, decreasing the value of a (moving from left to right) increases the magnification of the map, corresponding to an increased foveal representation. The original image on the left contains 512x512=256K pixels, as compared to the images after the complex log transform which contain 14,307 pixels, 5,000 pixels, and 496 pixels respectively (see (Rojer and Schwartz, 1990) for a complete discussion of the details involved in the design of a complex log transformation). Thus, algorithms can be expected run an order of magnitude or two faster in the log domain as compared to the Cartesian one in contemporary machine vision implementations.[2]

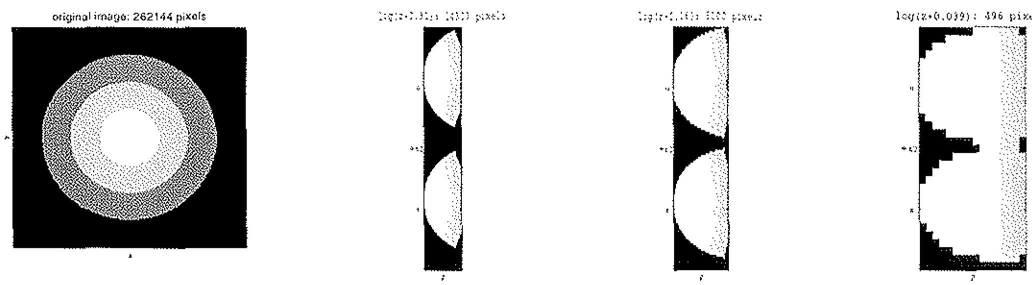


FIGURE 2.1. Example of an image (left), and its complex log transformation (right), for various values of the map parameter a. Note that decreasing a (moving from left to right) increases the representation of the foveal region in the log plane. Dark areas correspond to regions outside the domain of the mapping.

3. Diffusion in the log domain.

In this section we derive an algorithm for numerically integrating the anisotropic diffusion equation (1.2) directly in the log domain. This requires the derivation of the space-variant form of the gradient and divergence operators, the details which are given in appendix (A). For a more comprehensive examination of the form of a variety of differential operators in the complex log plane see (Fischl et al., 1996).

1. The complex log transformation requires a branch cut which divides the complex plane along the imaginary axis. This division was originally motivated by brain anatomy: the two half-planes in the range of the mapping correspond to the primary visual area in each hemisphere of the brain.
2. Rojer and Schwartz (Rojer and Schwartz, 1990) estimate that for biological systems the increase in speed can be up to four orders of magnitude.

Using the space variant forms of the gradient and the divergence (equations (A.8) and (A.11) respectively) we can write the anisotropic diffusion equation (1.2) in log coordinates as:

$$I_t = e^{-2\rho}\left((cI_\rho)_\rho + (cI_\phi)_\phi\right) \tag{3.1}$$

Where the ρ, ϕ and t subscripts denote partial differentiation with respect to the subscripted variable, and we have suppressed the arguments to $c()$ and $I()$ in the interests of conciseness. Substituting (3.1) into a Taylor series expansion of I around $t=t_0$ yields the first order approximation:

$$I(t_0+\Delta t) \approx I(t_0) + \Delta t\left(e^{-2\rho}(cI_\rho(t_0))_\rho + (cI_\phi(t_0))_\phi\right) \tag{3.2}$$

Using a discrete lattice with $\Delta\rho=\Delta\phi=1$, and considering the central pixel (ρ_0, ϕ_0), and its four connected neighbors (ρ_0, ϕ_{-1}), (ρ_0, ϕ_1), (ρ_{-1}, ϕ_0), and (ρ_1, ϕ_0) we use a centered difference approximation of the derivatives in (3.2). Labelling these pixels with superscripts 0, N, W, E, S respectively, we have:

$$\left(c^0(t_0)I^0{}_\rho(t_0)\right)_\rho \approx \frac{c^E(t_0)I^E{}_\rho(t_0) - c^W(t_0)I^W{}_\rho(t_0)}{2} \tag{3.3}$$

$$\left(c^0(t_0)I^0{}_\phi(t_0)\right)_\phi \approx \frac{\left(c^S(t_0)I^S{}_\phi(t_0) - c^N(t_0)I^N{}_\phi(t_0)\right)}{2} \tag{3.4}$$

We use both backwards and forward differences to approximate the partial derivatives with respect to the spatial variables so as to limit the domain of our numerical implementation to the four nearest neighbors of the central pixel:

$$\begin{aligned} I^W{}_\rho(t_0) &= I^0(t_0) - I^W(t_0),\ I^E{}_\rho(t_0) = I^E(t_0) - I^0(t_0), \\ I^N{}_\phi(t_0) &= I^0(t_0) - I^N(t_0),\ I^S{}_\phi(t_0) = I^S(t_0) - I^0(t_0) \end{aligned} \tag{3.5}$$

Substituting (3.3), (3.4) and (3.5) into (3.2) we arrive at:

$$I^0(t_0+\Delta t) \approx I^0(t_0)\left(1 - 0.5e^{-2\rho}\Delta t\left(\sum_{i\neq 0} c^i(t_0)\right)\right) + 0.5e^{-2\rho}\Delta t\left(\sum_{i\neq 0} c^i(t_0)I^i(t_0)\right) \tag{3.6}$$

Equation (3.6) can equivalently be written as the correlation of the image with a set of space and time varying masks:

$$I(\rho, \phi, t_0+\Delta t) \approx \sum_{\rho'}\sum_{\phi'} K^{t_0}{}_{\rho,\phi}(\rho', \phi')\, I(\rho+\rho', \phi+\phi', t_0) \tag{3.7}$$

Where the mask weights are given by:

$$K^{t_0}{}_{\rho,\phi} = \frac{e^{-2\rho}\Delta t}{2}\begin{bmatrix} 0 & c^N(t_0) & 0 \\ c^W(t_0) & \dfrac{2e^{2\rho}}{\Delta t} - \left(\displaystyle\sum_{i\neq 0} c^i(t_0)\right) & c^E(t_0) \\ 0 & c^S(t_0) & 0 \end{bmatrix} \tag{3.8}$$

In 2 dimensions the two components of the spatial gradient used in the computation of the conductance function are calculated using a Sobel operator with a negative exponential weight as specified by equation (A.8). At first sight equation (3.8) is distressing. It indicates that diffusion falls off exponentially with eccentricity. However, with the increased pixel spacing in the periphery comes increased numerical stability. An upper bound on the allowable stable time step Δt can be computed using Fourier-von Neumann stability analysis. In Cartesian space the numerical implementation will be stable if (Haberman, 1987):

$$\Delta t \le \frac{(\Delta x)^2}{4c} \tag{3.9}$$

If we choose c to be in the range (0,1] and let Δx=1, we then have Δt<=0.25 (the lower bound on c is necessary given the ill-posed nature of the backwards heat equation (Haberman, 1987, page 74). In the complex log plane the spatial grid of equation (3.9) is nonuniform. The interpixel distance is an exponential function of the radial coordinate, which implies that the stability constraint for an allowable time step in the log domain becomes

$$\Delta t \le \frac{e^{2\rho}}{4} \tag{3.10}$$

Equation (3.10) has important implications. It suggests that the nonlinear PDE (1.2) can be integrated using exponentially large time steps in the periphery, resulting in large scale structure enhancement in relatively few iterations. That is, we assume that t is approximately constant for a pixel and its four nearest neighbors, and allow the integration to proceed at different rates across the log domain image. Of course this is at the cost of fine scale peripheral image structure, but since such details aren't preserved in the periphery by the log mapping this is not a concern. Effectively, the space-variant time step allows different regions of the log plane to move through scale space at different rates - faster in the periphery and slower in the foveal region. Furthermore, if we replace Δt in equation (3.8) with $0.25e^{2\rho}$ the numerical implementation becomes identical to the implementation of the anisotropic diffusion equation in Cartesian space. That is, if we treat the log-domain image as a Cartesian one, then we are effectively allowing the integration to occur at different rates across the image.

4. Space-Variant Termination Conditions.

Given the variable-rate movement of the image through nonlinear scale space, a natural question to ask is how one determines a proper ending time for the numerical integration. If all points in the image are integrated for the same number of iterations, then different regions of the log image will represent different scales or times. If we wish to produce an image which is entirely at the same point in scale space, then the integration must be terminated in a space-variant manner. That is, we specify an ending time, then use equation (3.10) to determine whether a given ring of pixels (i.e. pixels of constant eccentricity) has reached the desired termination point, and if so, omit it from the domain of integration. In this way, the region of the image being integrated shrinks after each time step. The number of iterations required for each ring of pixels can be computed from equation (3.10) by fixing the desired number of iterations at some point in the log plane. Typically we fix the number of foveal iterations as this corresponds to the maximum number of time steps. Denoting the number of foveal time steps to be N_{fov}, we calculate the number of iterations as a function of the radial coordinate to be

$$N(\rho) = \frac{N_{fov} e^{2\log(a)}}{e^{2\rho}} \tag{4.1}$$

An example of this procedure is given in figures (4.1) and (4.2). Figure (4.1) depicts the num-

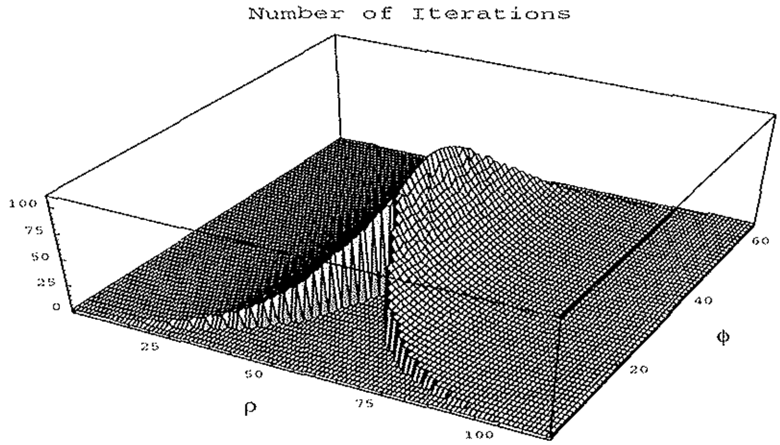


FIGURE 4.1. Number of iterations required as a function of log coordinate. Most of the periphery reaches the specified ending time after only a few iterations. while the fovea requires the full 100.

ber of iterations required as a function of log coordinate, with N_{fov}=100. As can be seen, the majority of the image reaches the specified termination point in less than 5 iterations, leaving only a small, shrinking foveal region to be integrated for the full 100 time steps. The original Cartesian image shown at the top left of figure (4.2) contains 580x720=417,600 pixels. The log image, shown in the top middle, is constructed by specifying the number of angular pixels (spokes) to be 64. Following (Rojer and Schwartz, 1990), this fixes the map parameter a=20.37 as well as the number of radial pixels (rings) to be 111. The total pixel count of the log image is therefore 7,104, a compression of more than 50 to 1. Given a and the size of the original image, the radial coordinate is constrained to be in the range $3.014 \le \rho \le 5.657$ which bounds the allowable time step through equation (3.10). Next, we fix the foveal integration time to be N_{fov}=100. In this case, the far periphery requires

$$N(\rho)\big|_{\rho=5.657} = \frac{100e^{2\log(20.37)}}{e^{2(5.657)}} \approx \frac{1}{2}\ iteration \tag{4.2}$$

Rounding up for numerical reasons, we can see that the far periphery arrives at the termination time in scale space in a single iteration! The efficacy of this scheme is illustrated in figure (4.2). The central image in the top row depicts the log mapping of the original image, while the bottom row shows the results of integrating the entire image for 1, 5, and 10 time steps from left to right, with a conductance function given by (Perona and Malik, 1987)

$$c_1(\nabla I) = e^{-\left(\frac{|\nabla I|}{k}\right)^2} \tag{4.3}$$

Where k is a real constant which controls the relationship between gradient magnitude and amount of diffusion.

Examining these images, we can see that peripheral details are quickly washed out while the fovea has yet to be enhanced. In contrast, the figure at the top right represents the variable end time diffusion we have outlined in this section. In this image, foveal noise has been smoothed while peripheral features are retained.

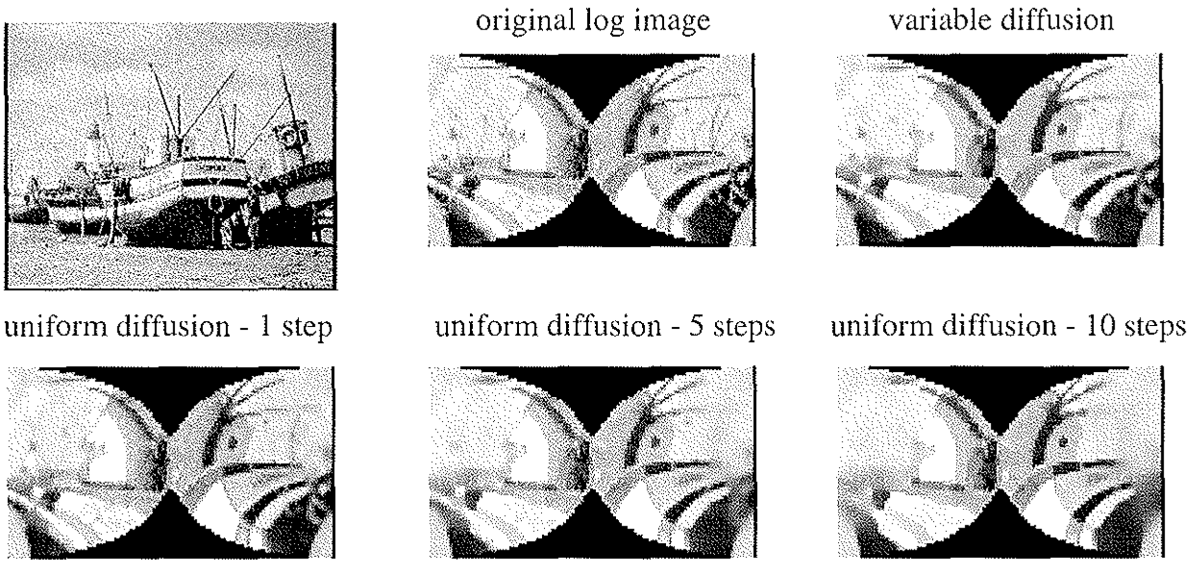


FIGURE 4.2. Types of integration termination schemes. Left: original noisy image. Top middle: log mapping of original image. Top right: variable termination diffusion. The far periphery terminates after only a single iteration, while the fovea integrates for the specified 50 steps. Bottom row: uniform integration for 1, 5, and 10 time steps from left to right.

We can quantify the computational savings provided by this scheme by examining the rate at which the domain of integration shrinks, as illustrated in figure (4.3). The left-hand plot shows the number of pixels in the domain of integration over time for the image shown in figure (4.2). The right-hand plot is the integral of the plot at the left, displaying the cumulative number of pixels integrated up to a given point in time as a fraction of the number of pixels in the full Cartesian image. Examining the left-hand plot, we can see that after only 4 iterations more than half of the image has reached the specified termination point, while by the 10th iteration the domain of integration has shrunk to less than a quarter of the image. The total number of pixels integrated over the full time span is approximately 79,000, or less than 1/5th of the number of pixels integrated in a single time step in the Cartesian domain, a computational cost decrease by a factor of 500. Another way to see the speed enhancement provided by the variable size time step is that integrating the image in this way requires the same amount of time as integrating the full log image for only 11 time steps. On a Sparc-20 the full Cartesian diffusion takes approximately 3,200 seconds, while the log diffusion requires a mere 8.4 seconds, or almost 400 times faster.

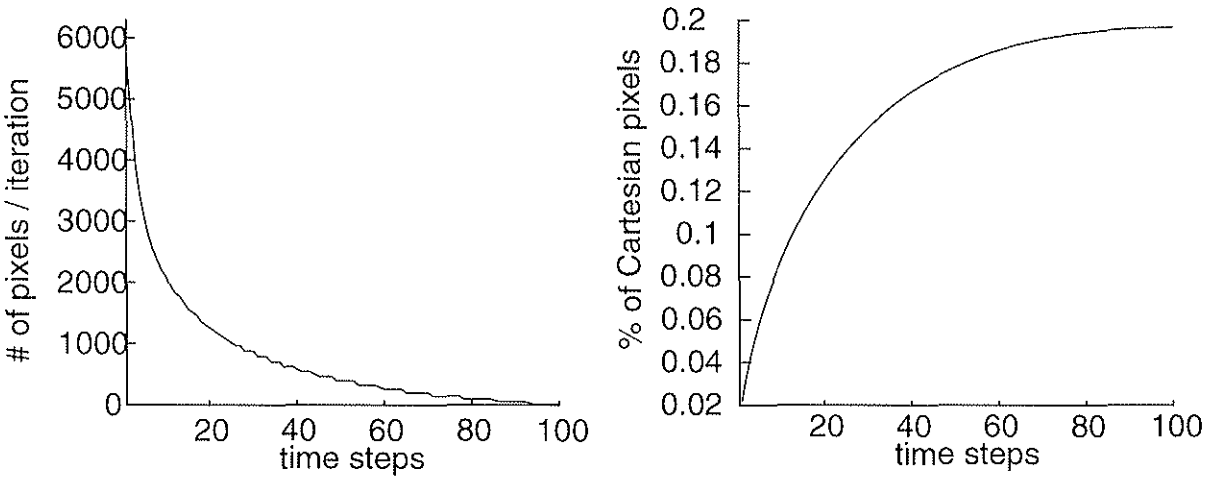


FIGURE 4.3. Number of pixels which require integration versus time. As time evolves, peripheral pixels arrive at the ending integration time and no longer require integration, thus reducing the effective pixel count of the image. The left hand plot shows the number of pixels being integrated at each time step. After 4 iterations approximately 1/2 of the log pixels still require integration. The right hand plot is the integral of the plot at the left, and illustrates the cumulative number of pixels which have been integrated at a given time as a fraction of the total pixel count of the Cartesian image. Integrating the log image for 100 foveal iterations corresponds to 1/5 of an iteration of the diffusion on the full Cartesian image.

5. Noise Tolerance.

An additional advantage provided by diffusion in the log domain is enhanced noise-tolerance Noise sensitivity has been shown to be problematic for the Perona and Malik diffusion using conductance functions such as equation (4.3) (Whitaker and Pizer, 1991; El-Fallah and Ford, 1994). This is due to the large gradient magnitudes arising in noisy image regions which inhibit diffusion, and are therefore preserved over time. Noise-tolerance is achieved naturally in the log plane diffusion due to the filtering which is necessary in the construction of the log image. Each log pixel has varying support in the Cartesian image, with the region of support growing with increasing eccentricity. The value of each log pixel is the average value of all Cartesian pixels which map to it.[1] In this way, moving *out* (i.e. increasing radial coordinate) in the log domain is equivalent to moving *up* (i.e. towards coarser scale images) in scale space. This process lowpass filters the image, effectively providing noise-tolerance for the diffusion in a manner similar to the multiscale approach suggested by (Whitaker and Pizer, 1991).

The enhanced noise-tolerance is illustrated in figure (5.1). From left to right these images are the original Cartesian image corrupted with white noise (0.1 amplitude), the image after undergoing Cartesian diffusion using the conductance function of equation (4.3) (100 iterations, k=0.05), the log mapping of the noisy image, and finally, the log image after diffusion using the same con-

1. The log plane images shown in this paper were all constructed using non-overlapping regions of support (Bederson et al., 1993). This approach can suffer from aliasing problems due to undersampling. A less efficient but more accurate means of constructing the log map is to use overlapping regions of support such that the image is sampled at or above the Nyquist rate at all eccentricities (Bonmassar and Schwartz, 1995).

ductance function (N_{fov}=100, k=0.2). The value of k used in the Cartesian diffusion in this image represents a compromise between two undesirable alternatives. Using this value of k, much of the license plate is enhanced at the cost of retaining shot noise in the left-hand side of the image, as well as the washing away of image detail such as the right-hand portion of the license plate. Setting k higher results in greater noise suppression at the cost of the destruction of more image structure; while a smaller value of k preserves more of the image, but also preserves more of the noise. In contrast, the log diffusion at the far right eliminates almost all of the noise while preserving most of the image detail contained in the original log image.

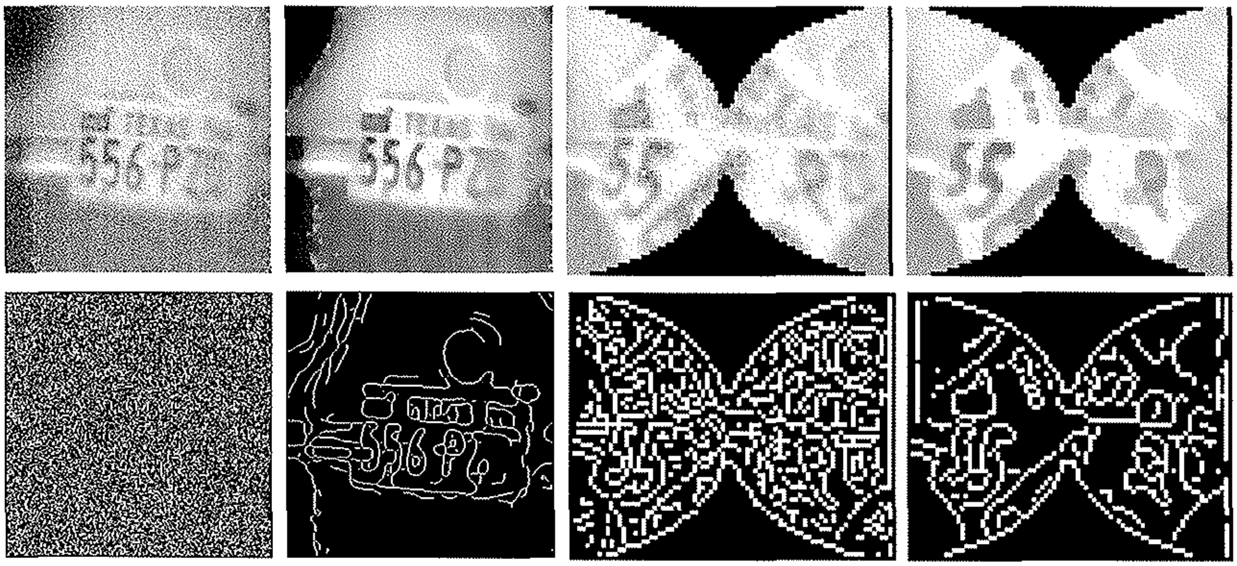


FIGURE 5.1. Noise tolerance of diffusion in the Cartesian and log domains. From left to right: original images, diffusion in Cartesian domain (100 iterations, k=0.05), complex log transform of the images, and anisotropic diffusion (N_{fov}=100 iterations, k=0.2) directly in the log plane. Bottom row: associated edge maps generated using a sobel operator followed by thresholding and non-maximum suppression.

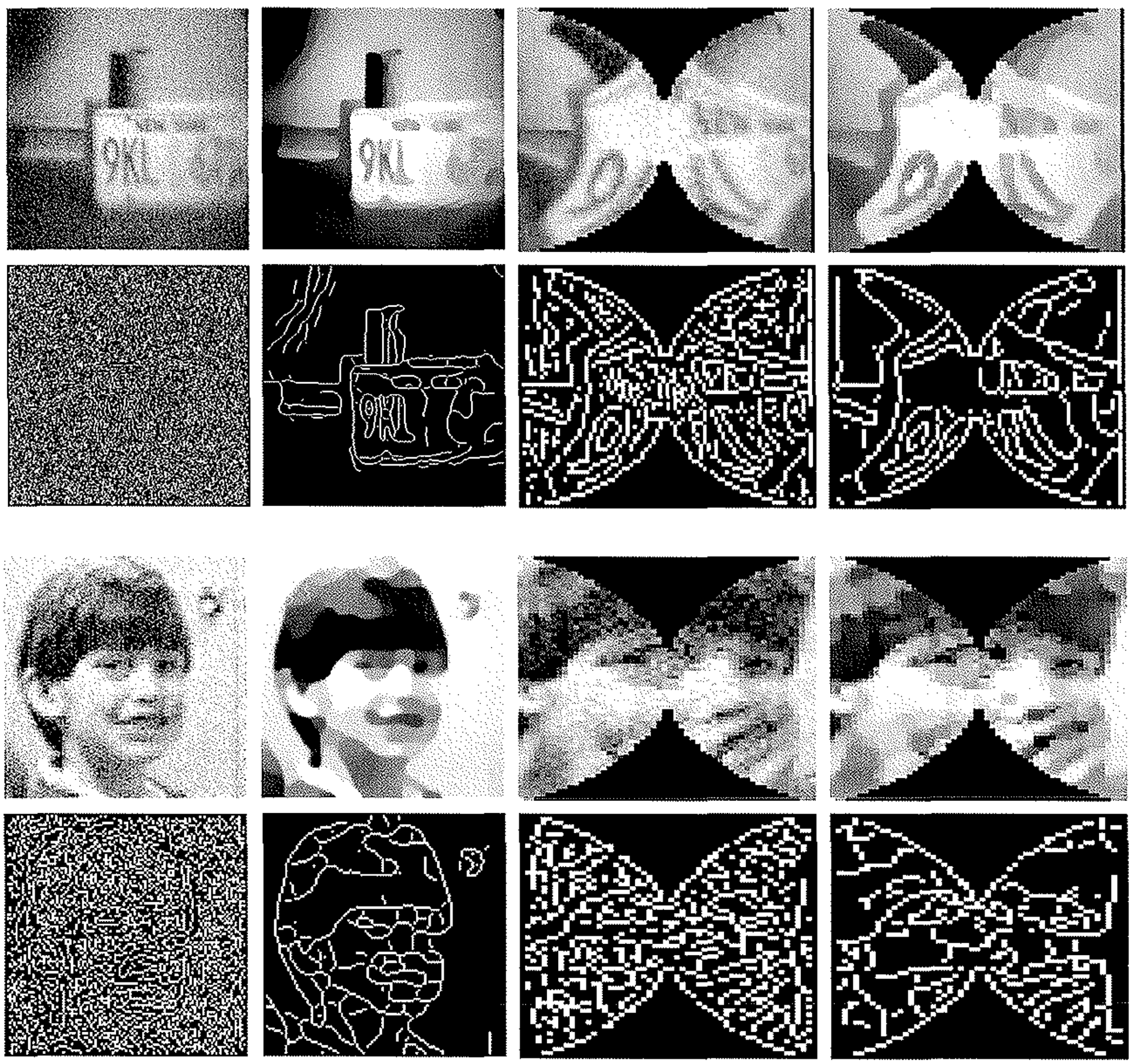
6. Results.

In the prior sections we showed that the form of a simple numerical implementation of the anisotropic diffusion equation in the log plane is equivalent to a variable grid size integration of the underlying PDE. In this section we show some results of applying equation (3.7) with the variable step size specified by (3.10) as well as the space-variant termination condition given by equation (4.1) on a variety of images. For comparison purposes, we also include the results of diffusion on the Cartesian images. Note that all images presented in this section have intensity values scaled to be in the range [0,1], and are integrated for the same number of effective time steps (100) using the same parameters (A=100, k=0.2). As noted earlier, the Perona and Malik scheme is incapable of dealing with the types of noise present in these images. We therefore use the conductance function proposed by El-Fallah and Ford for the Cartesian diffusion, as it has been shown to have good noise reduction characteristics (El-Fallah and Ford, 1994):

$$c_2(\nabla I) = \frac{1}{\sqrt{1 + A^2 |\nabla I|^2}} \tag{6.1}$$

Where the real constant A in this function plays a role similar to k in equation (4.3).

Figure (5.1) presents the results of Cartesian as well as log plane diffusion. From left to right the four columns in this figure are the original image, the Cartesian image after 100 time steps of anisotropic diffusion using the conductance function of equation (6.1), the complex log transformation of the image in column one, and the result of applying diffusion directly in the log plane with N_{fov}=100. Examining the log domain images, we can see that although much of the peripheral diffusion is accomplished in as few as 2-4 time steps, large scale structures such as the edge of the license plate and the boy's cheek are significantly enhanced in that time. To achieve comparable enhancement in the Cartesian domain requires between 50 and 100 iterations.



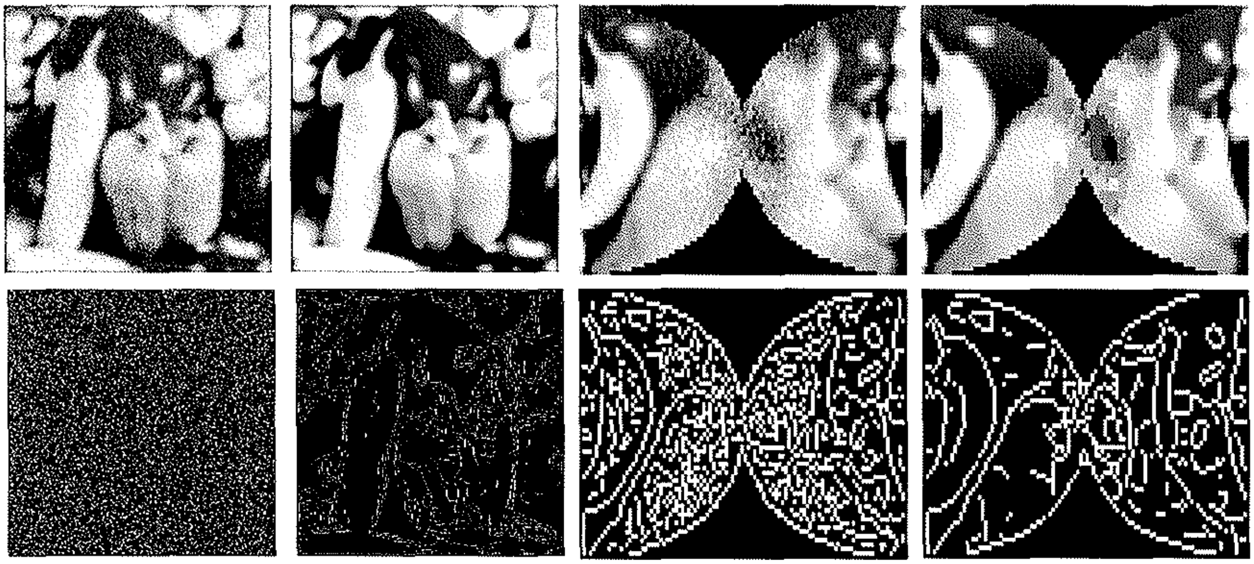


FIGURE 6.1. Anisotropic diffusion in the Cartesian and log domains. From left to right: original images, diffusion in Cartesian domain (100 iterations, A=100), complex log transform of the images, and anisotropic diffusion (N_{fov}=100 iterations, k=0.2) directly in the log plane. Edge maps are shown beneath each image for evaluation purposes.

7. Conclusion.

Diffusion is a powerful tool of great potential utility in machine as well as biological early vision systems (Gerrits and Vendrik, 1970; Cohen and Grossberg, 1984; Grossberg and Mingolla, 1985; Grossberg and Todorovic, 1988; Lee, 1995). It unifies multi-scale image enhancement and analysis into a simple procedure which yields a single image containing information at all scales of interest. Unfortunately, nonlinear diffusion is a computationally costly procedure, making it unsuitable for many real-time applications which might otherwise benefit from its use.

In this paper we have shown that anisotropic diffusion in conjunction with the complex log transformation has many desirable properties. First, the averaging that is a necessary part of the log mapping smooths noise in the original image, making the diffusion noise-tolerant when carried out in the log domain. More importantly, the non-uniform mesh spacing inherent in the log plane allows the use of integration rates which are exponential functions of eccentricity, yielding large scale enhancement in extremely few time steps.

From a biological standpoint, these results have important implications. If diffusion or a related process occurs at uniform *cortical* rates in mammalian visual cortex, then it is effectively proceeding at rates which are exponential functions of *retinal* eccentricity. Furthermore, coarse, features are enhanced prior to fine scale detail, providing large-scale contrast enhancement and noise reduction in as few as 2 or 3 time steps. While foveal diffusion remains problematic due to the severe constraints on the number of biologically plausible serial steps in rapid visual processing, at great enough eccentricity the exponentially increasing integration rate yields enhancement in few enough time steps to make peripheral and possibly perifoveal diffusion a biologically plausible process.

From a machine vision standpoint, the performance increase obtained due to the combination of the compressive effects of the log mapping as well as the exponential integration rates is between two and three orders of magnitude. In conjunction with current DSP architectures, this type of speed increase makes the use of anisotropic diffusion for image enhancement realistically available for real-time vision applications.

8. Bibliography.

Alvarez, L., Lions, P.-L., and Morel, J.-M. (1992). Image selective smoothing and edge detection by nonlinear diffusion. II. *SIAM Journal of Numerical Analysis*, 29(3):845–866.

Alvarez, L. and Mazorra, L. (1994). Signal and image restoration using shock filters and anisotropic diffusion. *SIAM Journal of Numerical Analysis*, 31(2):590–605.

Bederson, B. B., Wallace, R. S., and Schwartz, E. L. (1993). A miniaturized space-variant active vision system: Cortex-I. In Mammone, R. J., editor, *Artificial Neural Networks for Speech and Vision*, chapter 24, pages 429–455. Chapman & Hall, London.

Bonmassar, G. and Schwartz, E. (1995). Space-variant fourier analysis: the exponential chirp transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, submitted.

Bonmassar, G. and Schwartz, E. (1996a). Fourier analysis and cortical architectures: the exponential chirp transform. *Real Time Imaging*, In press.

Bonmassar, G. and Schwartz, E. L. (1994). Geometric invariance in space-variant vision systems: the exponential chirp transform. In *ICPR Proceedings*, ICPR-12, pages 204–207. International Conference on Pattern Recognition.

Bonmassar, G. and Schwartz, E. L. (1996b). Lie groups, space-variant fourier analysis and the exponential chirp transform. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 18-20, 1996, San Francisco CA, USA. 492-498.*

Campbell, F. and Green, D. (1965). Optical quality of the human eye. *Journal of Physiology, London*, 186:558–578.

Catte, F., Lions, P.-L., Morel, J.-M., and Coll, T. (1992). Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal of Numerical Analysis*, 29(1):182–193.

Churchill, R. V. and Brown, J. W. (1984). *Complex variables and applications*. McGraw-Hill Book Company, New York.

Cohen, M. A. and Grossberg, S. (1984). Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance. *Perception and Psychophysics*, 36:428–456.

Cottet, G.-H. and Germain, L. (1993). Image processing through reaction combined with nonlinear diffusion. *Mathematics of Computation*, 61(204):659–673.

Dang, T., Olivier, J., and Maitre, H. (1994). An image segmentation technique based on edge preserving smoothing filter and anisotropic diffusion. In *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation*, pages 65–69.

Daniel, P. M. and Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. *Journal of Physiology*, 159:203–221.

El-Fallah, A. I. and Ford, G. E. (1994). Nonlinear adaptive image filtering based on inhomogeneous diffusion and differential geometry. *SPIE Image and Video Processing II*, 2182:49–63.

Engquist, B., Lotstedt, P., and Sjogreen, B. (1989). Nonlinear filters for efficient shock computation. *Mathematics of Computation*, 52(186):509–537.

Fischl, B., Cohen, M. A., and Schwartz, E. L. (1996). The local structure of space-variant images. Technical Report CAS/CNS-TR-95-023, Boston University, Department of Cognitive and Neural Systems, 677 Beacon St., Boston MA.

Fischl, B. and Schwartz, E. (1995). Learning an integral equation approximation to nonlinear anisotropic diffusion in image processing. Technical Report CAS/CNS-TR-95-033, Boston University, Department of Cognitive and Neural Systems, 677 Beacon St., Boston MA.

Fischl, B. and Schwartz, E. L. (1996). Learned adaptive nonlinear filtering for anisotropic diffusion approximation in image processing. In *International Conference on Pattern Recognition*.

Gerrits, H. and Vendrik, J. (1970). Simultaneous contrast, filling-in process and information processing in man's visual system. *Experimental Brain Research*, 11:411–430.

Grossberg, S. and Mingolla, E. (1985). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, 92(2):173–211.

Grossberg, S. and Todorovic, D. (1988). Neural dynamics of 1-d and 2-d brightness perception: A unified model of classical and recent phenomena. *Perception and Psychophysics*, 43:241–277.

Haberman, R. (1987). *Elementary applied partial differential equations*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, second edition.

Hummel, A. (1986). Representations based on zero-crossings in scale-space. In Fischler, M. and Firschein., O., editors, *Readings in Computer Vision: Issues, Problems, Principles and Paradigms*. Morgan Kaufmann, Los Angeles.

Illner, R. and Neunzert, H. (1993). Relative entropy maximization and directed diffusion equations. *Mathematical Methods in the Applied Sciences*, 17:545–554.

Kacur, J. and Mikula, K. (1995). Solution of nonlinear diffusion appearing in image smoothing and edge detection. *Applied Numerical Mathematics*, 50:47–59.

Koenderink, J. (1984). The structure of images. *Biological Cybernetics*, 50:363–370.

Lee, T. S. (1995). A bayesian framework for understanding texture segmentation in the primary visual cortex. *Vision Research*, 35(18):2643.

Li, X. and Chen, T. (1994). Nonlinear diffusion with multiple edginess thresholds. *Pattern Recognition*, 27(8):1029–1037.

Malladi, R. and Sethian, J. A. (1995). Image processing via level set curvature flow. *Proceedings of the National Academy of Sciences*, 92:7046–7050.

Marr, D. and Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London*, B 207:187–217.

Merigan, W. H. and Katz, L. M. (1990). Spatial resolution across the macaque retina. *Vision Research*, 30(7):985–991.

Messner, R. and Szu, H. (1986). An image processing architecture for real time generation of scale and rotation invariant patterns. *Computer Vision, Graphics and Image Processing*, 31:50–66.

Nitzberg, M. and Shiota, T. (1992). Nonlinear image filtering with edge and corner enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(8):826–833.

Nördstrom, N. K. (1990). Biased anisotropic diffusion: a unified regularization and diffusion approach to edge detection. *Image and Vision Computing*, 8(4):318–327.

Oram, M. W. and Perrett, D. I. (1992). Time course of neural responses discriminating different views of the face and head. *Journal of Neurophysiology*, 68(1):70–84.

Osher, S. and Rudin, L. I. (1990). Feature-oriented image enhancement using shock filters. *SIAM Journal of Numerical Analysis*, 27(4):919–940.

Pauwels, E. J., Proesmans, M., Gool, L. J. V., Moons, T., and Oosterlinck, A. (1993). Segmentation and image enhancement using coupled anisotropic diffusion. *SPIE*, 2094:836–847.

Perona, P. and Malik, J. (1987). Scale space and edge detection using anisotropic diffusion. In *Proceedings of the IEEE Computer Society Workshop on Computer Vision*, pages 16–27, Miami, FL.

Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639.

Perona, P., Shiota, T., and Malik, J. (1994). Anisotropic diffusion. In Romeny, B. M. T. H., editor, *Geometry Driven Diffusion in Computer Vision*, chapter 3, pages 73–92. Kluwer.

Price, C. B., Wambacq, P., and Oosterlinck, A. (1990). Image enhancement and analysis with reaction-diffusion paradigm. *IEE Proceedings*, 137 Pt. I(3):136–145.

Rojer, A. S. and Schwartz, E. L. (1990). Design considerations for a space-variant visual sensor with complex-logarithmic geometry. In *10th International Conference on Pattern Recognition*, volume 2, pages 278–285.

Sandini, G., Bosero, F., Bontino, F., and Ceccherini, A. (1989). The use of an anthropomorphic visual sensor for motion estimation and object tracking. *Proceedings of OSA optical meeting on image understanding and machine vision.*

Sandini, G. and Dario, P. (1989). Active vision based on space-variant sensing. *International symposium on robotics research.*

Schenker, P., Cande, E., Wong, K., and Patternson, W. (1981). New sensor geometries for image processing: computer vision in the polar exponential grid. *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing.*, pages 1144–1148.

Schwartz, E. (1994). Computational studies of the spatial architecture of primate visual cortex: Columns, maps, and protomaps. In Peters, A. and Rocklund, K., editors, *Primary visual cortex in primates*, volume 10 of *Cerebral Cortex*, chapter 9. Plenum Press, New York.

Schwartz, E. L. (1977). Spatial mapping in the primate sensory projection: analytic structure and relevance to perception. *Biological Cybernetics*, 25:181–194.

Schwartz, E. L. (1980). Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding. *Vision Research*, 20:645–669.

Shah, J. (1996). A common framework for curve evolution, segmentation and anisotropic diffusion. In *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 136–142, San Francisco, CA.

Thorpe, S. J. and Imbert, M. (1989). Biological constraints on connectionist models. In Pfeifer, R., Schreter, Z., and Fogelman-Soulie, F., editors, *Connectionism in perspective*, pages 63–92. Elsevier, Amsterdam.

Vogels, R. and Orban, G. A. (1991). Quantitative study of striate single unit responses in monkey performing an orientation task. *Experimental Brain Research*, 84:1–11.

Wässle, H., Grünert, U., Röhrenbeck, J., and Boycott, B. B. (1990). Retinal ganglion cell density and cortical magnification factor in the primate. *Vision Research*, 30(11):1897–1911.

Weiman, C. (1988). 3-d sensing with polar exponential sensory arrays. *SPIE Conference on digital and optical shape representation and pattern recognition.*

Wertheim, T. (1894). Uber die indirekte sehscharfe. *Zeitschrift fur Psychologie und Physiologie der Sinnesorgane*, 7:172–187.

Whitaker, R. T. (1993). Geometry-limited diffusion in the characterization of geometric patches in images. *CVGIP: Image Understanding.*, 57(1):111–120.

Whitaker, R. T. and Pizer, S. M. (1991). A multi-scale approach to nonuniform diffusion. *Computer Vision, Graphics and Image Processing.*, 57:99–110.

Williams, D. and Coletta, N. (1987). Cone spacing and the visual resolution limit. *Journal of the Optical Society of America*, 4:1514–1523.

Wilson, H. R., Levi, D., Maffei, L., Rovamo, J., and DeValois, R. (1990). The perception of form: retina to striate cortex. In Spillmann, L. and Werner, J. S., editors, *Visual Perception: The Neurophysiological Foundations*, chapter 10, pages 231–272. Academic Press.

Witken, A. (1983). Scale-space filtering. In *International Joint Conference on Artificial Intelligence*, pages 1019–1021. Karlsruhe, West Germany.

Yamamoto, H., Yeshurun, Y., and Levine, M. D. (1996). An active foveated vision system: attentional mechanisms and scan path convergence measures. *Computer vision and image understanding: CVIU*, 63(1):50.

9. Appendices.

A. Differential operators.

In this section we compute the form of the ∇ operator, which yields the space-variant form of the gradient and the divergence, As noted in section (2), the complex log coordinate transform considered in this work is of the form:

$$w = \log(z+a), a \in \Re, \quad z, w \in C, \ Re(z) \geq 0 \tag{A.1}$$

More explicitly, the log coordinates (ρ,ϕ) are given in terms of their Cartesian counterparts (x,y) by:

$$\rho = \log\left(\sqrt{(x+a)^2 + y^2}\right), \ \phi = \tan^{-1}\left(\frac{y}{(x+a)}\right) \tag{A.2}$$

The inverse relations are:

$$x = e^{\rho}\cos\phi - a, \ y = e^{\rho}\sin\phi \tag{A.3}$$

The log mapping of equation (A.2) as well as the inverse mapping given by (A.3) are both analytic everywhere in their respective domains, and are hence conformal. This has a number of interesting and useful implications. For the present purposes, the most important of these is that the conformal nature of the mapping ensures that local angles are preserved (Churchill and Brown, 1984). This in turn implies that the log-polar coordinate basis is orthogonal when projected into Cartesian space. This fact will be used to simplify the derivation of the log domain gradient in section (A.1).

A.1. Space-Variant form of ∇f.

The conformal nature of the complex log mapping yields a simple derivation of the form of the gradient in the log domain. As noted in the introduction to this section, the conformality of the log mapping implies that local angles are preserved by the transformation. This simplifies the derivation considerably. Specifically, it insures that the basis vectors of the (ρ,ϕ) space which are orthogonal in the log domain, are also orthogonal when projected into Cartesian space (see figure (A.4)). Since the gradient is the combination of the directional derivative in *any* two orthogonal directions, we are assured that the gradient in the log space is of the form

$$\nabla f = A(\rho, \phi)\left(\frac{\partial f}{\partial \rho}\boldsymbol{e}_{\rho} + \frac{\partial f}{\partial \phi}\boldsymbol{e}_{\phi}\right) \qquad \text{(A.4)}$$

Where $\boldsymbol{e}_{\rho}$ and $\boldsymbol{e}_{\phi}$ are an orthonormal basis (in the induced metric) for the log domain, and the $A(\rho,\phi)$ term accounts for the variation in length a vector experiences under the log mapping. Note that equation (A.4) holds for *any* conformal mapping, with the specifics of the transformation expressed in the coefficient function A. Another way to see that the gradient must be of the form given in (A.4) is to observe that any inhomogenous scaling of the basis vectors would result in the angle between the gradient and the basis vectors being different in the two spaces, which cannot be the case since the mapping is conformal. All that remains is to determine the form of the coefficient function. To do so, we use the invariance of the magnitude of the gradient under a change of coordinates. That is, the length of the gradient (or its square) must be the same in both domains. Hence:

$$A^2\left(f_{\rho}^2 + f_{\phi}^2\right) = f_x^2 + f_y^2 \qquad \text{(A.5)}$$

Using the chain rule to express $\partial f/\partial \rho$ and $\partial f/\partial \phi$ in terms of f_x and f_y yields:

$$A^2\left((f_x x_{\rho} + f_y y_{\rho})^2 + (f_x x_{\phi} + f_y y_{\phi})^2 \right) = f_x^2 + f_y^2 \qquad \text{(A.6)}$$

Expanding (A.6) using the derivatives of equation (A.3) and solving for A results in

$$A(\rho, \phi)^2\left(\left(f_x e^{\rho} c + f_y e^{\rho} s\right)^2 + \left(-f_x e^{\rho} s + f_y e^{\rho} c\right)^2\right) = f_x^2 + f_y^2 \qquad \text{(A.7a)}$$

$$A(\rho, \phi)^2\left(f_x^2 e^{2\rho} c^2 + f_y^2 e^{2\rho} s^2 + 2 f_y f_x e^{2\rho} sc + f_x^2 e^{2\rho} c^2 + f_y^2 e^{2\rho} s^2 - 2 f_y f_x e^{2\rho} sc\right) = f_x^2 + f_y^2 \qquad \text{(A.7b)}$$

$$\Rightarrow A(\rho, \phi) = e^{-\rho} \qquad \text{(A.7c)}$$

Where $s = \sin\phi$ and $c = \cos\phi$. Thus, the gradient in the space-variant domain is given by[1]:

$$\nabla f = e^{-\rho}\left(\frac{\partial f}{\partial \rho}\boldsymbol{e}_{\rho} + \frac{\partial f}{\partial \phi}\boldsymbol{e}_{\phi}\right) \qquad \text{(A.8)}$$

From equation (A.8) it is apparent that the ∇ operator has the general form $e^{-\rho}(\partial/\partial\rho\ \boldsymbol{e}_{\rho} + \partial/\partial\phi\ \boldsymbol{e}_{\phi})$ which allows the direct computation of quantities such as the divergence and the curl in the log plane.

---

1. Note that this derivation does not account for the varying support of each log pixel. As one moves into the periphery of the log plane, each log pixel is typically generated by averaging a larger region of Cartesian space, both in the mammalian retina and in machine vision systems. The averaging is done to avoid aliasing in the periphery, and attenuates high frequency information, partially offsetting the need for a negative exponential weighting to account for varying pixel separation.

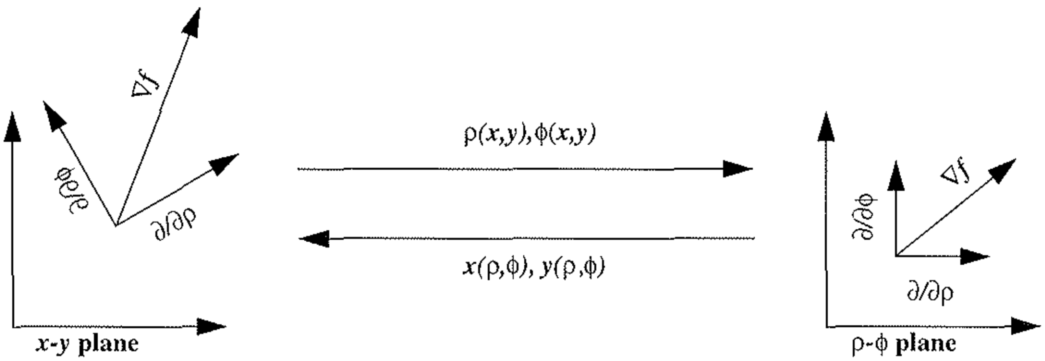


FIGURE A.4. Representation of the relationship between the basis vectors and the gradient in the two spaces. The mapping preserves the angles between the vectors, but not their lengths.

A.2. Space-Variant form of $\nabla \cdot f$.

The form of the divergence of a vector field in the log plane can be calculated in a straightforward manner using the form of the ∇ operator derived in the prior section. To do so we will require the derivatives of the log plane orthonormal basis vectors $\boldsymbol{e}_\rho$ and $\boldsymbol{e}_\phi$ with respect to the log coordinates. Like their polar counterparts, $\boldsymbol{e}_\rho$ and $\boldsymbol{e}_\phi$ do not change in the radial direction and hence both derivatives with respect to ρ are zero. To calculate the change in the basis vector with respect to the angular log coordinate we use the chain rule as follows:

$$\boldsymbol{e}_\rho = \cos\phi\frac{\partial}{\partial x} + \sin\phi\frac{\partial}{\partial y},\ \boldsymbol{e}_\phi = \cos\phi\frac{\partial}{\partial y} - \sin\phi\frac{\partial}{\partial x} \tag{A.9a}$$

$$\frac{\partial \boldsymbol{e}_\rho}{\partial \phi} = \cos\phi\frac{\partial}{\partial y} - \sin\phi\frac{\partial}{\partial x} = \boldsymbol{e}_\phi,\ \frac{\partial \boldsymbol{e}_\phi}{\partial \phi} = -\sin\phi\frac{\partial}{\partial x} - \cos\phi\frac{\partial}{\partial y} = -\boldsymbol{e}_\rho \tag{A.9b}$$

Given these relations, the divergence of an arbitrary vector field whose components expressed in the orthonormal log basis ($\boldsymbol{e}_\rho$, $\boldsymbol{e}_\phi$) are (f^ρ, f^ϕ) can be calculated as:

$$\nabla \cdot f = e^{-\rho}\left(\frac{\partial}{\partial \rho}\boldsymbol{e}_\rho + \frac{\partial}{\partial \phi}\boldsymbol{e}_\phi\right) \cdot \left(f^\rho \boldsymbol{e}_\rho + f^\phi \boldsymbol{e}_\phi\right) \tag{A.10}$$

Using equations (A.9b) and the orthonormality of the basis vectors, the divergence simplifies to:

$$\nabla \cdot f = e^{-\rho}\left(f^\rho_\rho + f^\phi_\phi + f^\rho\right) \tag{A.11}$$