

GRADE: Gibbs Reaction And Diffusion Equations

Song Chun Zhu

David Mumford

Computer Science Department
Stanford University
Stanford, CA 94305

Division of Applied Mathematics
Brown University
Providence, RI 02912

Abstract

Recently there have been increasing interests in using nonlinear PDEs for applications in computer vision and image processing, in this paper, we propose a general statistical framework for designing a new class of PDEs. For a given application, a Markov random field model $p(\mathbf{I})$ is learned according to the minimax entropy principle studied in [25][26], so that $p(\mathbf{I})$ should characterize the ensemble of images in our application. $p(\mathbf{I})$ is a Gibbs distribution whose energy terms can be divided into two categories. Subsequently the partial differential equations given by gradient descent on the Gibbs potential are essentially reaction-diffusion equations, where the energy terms in one category produce anisotropic diffusion while the **inverted** energy terms in the second category produce reaction associated with pattern formation. We call this new class of PDEs the Gibbs Reaction And Diffusion Equations—GRADE and we demonstrate experiments where GRADE are used for texture pattern formation, denoising, image enhancement, and clutter removal.

1 Introduction

Nonlinear partial differential equations (PDEs) have long been adopted in modeling a vast variety of phenomena in chemistry, physics, and biology. Recently some of these PDEs have inspired interesting methods for solving problems in computer vision, image processing, and graphics, among which the following two examples are remarkable. (I) The Turing reaction-diffusion equations for texture pattern formation [21][22][23], and its application in image processing [18][19]. (II) The Perona-Malik anisotropic diffusion equations for generating scale space and for noise removal [17][16].

Despite the success of these PDEs, there lacks a general theory to guide the design of PDE in vision applications. As the images in a given application are often not directly related to the chemical or physical processes from which the PDEs are originally derived,

there is little justification for applying these PDEs in computer vision.

This paper proposes a statistical framework for designing a new class of PDE, inspired by our previous work on modeling texture and natural images. In two previous papers [25][26], the authors have been studying a minimax entropy theory based on which a Gibbs distribution $p(\mathbf{I})$ is learned to characterize a set of observed images. We find that the learned potential functions in $p(\mathbf{I})$ can be divided into two classes, and subsequently the partial differential equation given by gradient descent on the Gibbs potential is essentially reaction-diffusion equation, which we named the *Gibbs Reaction And Diffusion Equation—GRADE*. The GRADE consists of two components: the energy terms from one class produce anisotropic diffusion, while the energy terms in the second class produce reaction associated with pattern formation and enhancing preferred image features. Unlike the PDEs derived from a chemical or physical process, The GRADE is well learned from a given application, and their behaviors are specified by the Gibbs probability distributions. We demonstrate experiments where the GRADE is used for texture pattern formation, denoising, image enhancement, and clutter removal.

This paper is organized as follows. Section (2) briefly reviews previous work on nonlinear PDEs. Section (3) derives the Gibbs reaction and diffusion equation and section (4) analyzes its properties. Section (5) presents experiments on pattern synthesis, denoising, image enhancement, and clutter removal. Then section (6) concludes with a discussion.

2 Previous work

2.1 Reaction-diffusion for pattern formation

A set of nonlinear PDEs was first studied in [21] for modeling the formation of animal coat patterns by the diffusion and reaction of chemicals, which Turing called the “morphogens”. These equations were fur-

ther explored by Murray in theoretical biology [14]. For example, let $A(x, y, t)$ and $B(x, y, t)$ be the concentrations of two morphogens at location (x, y) and time t , the typical reaction-diffusion equations are,

$$\begin{cases} \frac{\partial A}{\partial t} = D_a \Delta A + R_a(A, B) \\ \frac{\partial B}{\partial t} = D_b \Delta B + R_b(A, B) \end{cases}$$

where D_a, D_b are constants, $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the Laplacian operator, and $R_a(A, B), R_b(A, B)$ are nonlinear functions for the reaction of chemicals, e.g., $R_a(A, B) = A * B - A - 12$ and $R_b(A, B) = 16 - A * B$.

The morphogen theory provides a way for synthesizing some texture patterns [22][23]. In the texture synthesis experiments, chemical concentrations are replaced by various colors, and the equations run for a finite steps with free boundary condition starting with some initial patterns. In some cases, both the initial patterns and the running processes have to be controlled manually in order to generate realistic textures. Two canonical textures synthesized by the Turing reaction-diffusion equation are the leopard blobs and zebra stripes. Recently these equations are utilized in image processing such as enhancement of fingerprint images and image halftoning [18][19].

In the reaction-diffusion equations above, the reaction terms are responsible for pattern formation, however they also make the equations unstable or unbounded. Even for a small system, the existence and stability problems for these PDEs are intractable (Grindrod 1996). In fact, we believe that running any nonlinear PDEs for a finite steps will render some patterns, but it is unknown how to design a set of PDEs for a given texture pattern.

2.2 Anisotropic diffusion

Perona and Malik introduced a family of anisotropic diffusion equations for generating image scale space $\mathbf{I}(x, y, t)$ [17]. Their equation simulate the "heat" diffusion process,

$$\mathbf{I}_t = \text{div}(c(x, y, t) \nabla \mathbf{I}), \quad \mathbf{I}(x, y, 0) = \mathbf{I}_0, \quad (1)$$

where $\nabla \mathbf{I} = (\mathbf{I}_x, \mathbf{I}_y)$ is the intensity gradient and div is the divergence operator, $\text{div}(\vec{V}) = \nabla_x P + \nabla_y Q$ for any vector $\vec{V} = (P, Q)$. In practice, the heat conductivity $c(x, y, t)$ is defined as function of location gradients, for example, Perona and Malik chose

$$\mathbf{I}_t = \nabla_x \left(\frac{1}{1 + (\mathbf{I}_x/b)^2} \mathbf{I}_x \right) + \nabla_y \left(\frac{1}{1 + (\mathbf{I}_y/b)^2} \mathbf{I}_y \right), \quad (2)$$

where b is a scaling constant. It is easy to see that equation (2) minimizes the following energy function

by gradient descent,

$$U(\mathbf{I}) = \int \int \lambda(\nabla_x \mathbf{I}) + \lambda(\nabla_y \mathbf{I}) dx dy,$$

where $\lambda(\xi) = \text{const.} \log(1 + (\xi/b)^2)$ and $\lambda'(\xi) = \text{const.} \frac{\xi}{1 + (\xi/b)^2}$ are plotted in figure 1. Similar forms of the energy are widely used as prior distributions in image restoration and segmentation[6][8][13][4].

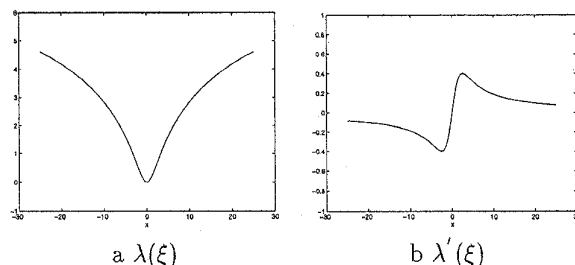


Figure 1: a. $\lambda(\xi) = \text{const.} \log(1 + (\xi/b)^2)$, b. $\lambda'(\xi) = \text{const.} \frac{\xi}{1 + (\xi/b)^2}$

Although the anisotropic diffusion equations can be adopted for removing noise and enhancing edges [16], $\mathbf{I}(x, y, t)$ converges to a flat image as $t \rightarrow \infty$ in the Perona-Malik equation. A robust anisotropic diffusion equation is recently studied in [1].

3 From Gibbs distributions to reaction-diffusion equations

Let \mathbf{I} be an image defined on an $N \times N$ lattice L . In previous work on modeling textures and real world scenes, we proposed a new class of Gibbs distributions of the following form [25][26],

$$p(\mathbf{I}; \Lambda, S) = \frac{1}{Z} e^{-U(\mathbf{I}; \Lambda, S)}$$

In the above equation, Z is a normalization constant, $S = \{F_1, F_2, \dots, F_n\}$ is a set of filters to characterize the essential features of the images, and $\Lambda = \{\lambda_1(), \dots, \lambda_n()\}$ a set of potential functions on the features extracted by S . The potential is

$$U(\mathbf{I}; \Lambda, S) = \sum_{i=1}^n \sum_{(x, y) \in L} \lambda_i(F_i * \mathbf{I}_{(x, y)}), \quad (3)$$

where $F_i * \mathbf{I}_{(x, y)}$ is the filter response of F_i at (x, y) .

In practice, S is chosen by minimizing the entropy of $p(\mathbf{I})$ from a bank of filters such as the Gabor filters at various scales and orientations [3][2] and wavelet transforms [12][20]. In general, these filters can be

non-linear functions of the image intensities. In the rest of this paper, we shall only study the following linear filters.

1. An intensity filter $\delta()$, and gradient filters ∇_x, ∇_y .
2. The Laplacian of Gaussian filters.

$$LG(x, y, s) = \text{const} \cdot (x^2 + y^2 - s^2) e^{-\frac{x^2+y^2}{s^2}}, \quad (4)$$

where $s = \sqrt{2}\sigma$ stands for the scale of the filter. We denote these filters by $LG(s)$. A special filter is $LG(\frac{\sqrt{2}}{2})$, which has a 3×3 window $[0, \frac{1}{4}, 0; \frac{1}{4}, -1, \frac{1}{4}; 0, \frac{1}{4}, 0]$, and we denote it by Δ .

3. Gabor filters with both sine and cosine components.

$$G(x, y, s, \theta) = \text{const} \circ \text{Rot}(\theta) \circ e^{-\frac{1}{2s^2}(4x^2+y^2)} e^{-i\frac{2\pi}{s}x}. \quad (5)$$

It is a sine wave at frequency $\frac{2\pi}{s}$ modulated by an elongated Gaussian function, and rotated at angle θ . We denote the real and image parts of $G(x, y, s, \theta)$ by $G\cos(s, \theta)$ and $G\sin(s, \theta)$.

In equation (3), the filter responses are quantitized into a finite number of bins and $\lambda_i(), i = 1, 2, \dots, n$ becomes a piecewisely constant function or a vector. Thus the form of $\lambda_i(), i = 1, 2, \dots, n$ is learned from a set of training images in a fully non-parametric way so that $p(\mathbf{I})$ can reproduce the marginal distributions of $F_i * \mathbf{I}^{obs}, i = 1, 2, \dots, n$ over a set of training images \mathbf{I}^{obs} in a given application.

Verified by stochastic sampling, the learned distribution is capable of modeling a wild variety of images. A detailed account of this theory is referred to [25].

In contrast to existing models on natural images and texture patterns in the literature, we found that the learned $\lambda_i(), i = 1, 2, \dots, n$ can be divided into two classes. For example, two typical $\lambda_i()$ s are plotted in figure 2, which are fit to a family of functions (see the dashed curves),

$$\begin{aligned} \phi(\xi) &= a \left(1 - \frac{1}{1 + (|\xi - \xi_0|/b)^\gamma} \right) \quad a > 0 \\ \psi(\xi) &= a \left(1 - \frac{1}{1 + (|\xi - \xi_0|/b)^\gamma} \right) \quad a < 0, \end{aligned}$$

where ξ_0, b are respectively the translation and scaling constants, and $|a|$ weights the contribution of the filter. $\phi(\xi)$ assigns lowest energy (or highest probability) to filter responses around ξ_0 and $\xi_0 = 0$ in most cases, and $\psi(\xi)$ has lower energy at the two tails which represent salient image features such as object boundaries. These *inverted* potential functions are in contrast to all previous image models, and they are essential for modeling realistic images.

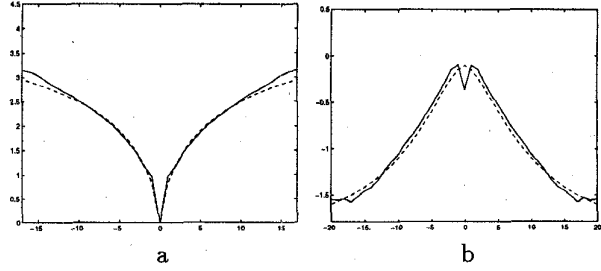


Figure 2: Learned potential functions $\lambda_i()$. The dashed curves are fitting functions: a). fitting $\phi(\xi)$ with $(a = 5, b = 10, \xi_0 = 0, \gamma = 0.7)$. b). Fitting $\psi(\xi)$ with $(a = -2.0, b = 10, \xi_0 = 0, \gamma = 1.6)$.

Thus we write equation (3) in the following form,

$$U(\mathbf{I}; \Lambda, S) = \sum_{i=1}^{n_d} \sum_{(x,y)} \phi_i(F_i * \mathbf{I}(x,y)) + \sum_{i=n_d+1}^n \sum_{(x,y)} \psi_i(F_i * \mathbf{I}(x,y)). \quad (6)$$

Note that the filter set is divided into two parts $S = S_d \cup S_r$, with $S_d = \{F_i, i = 1, 2, \dots, n_d\}$ and $S_r = \{F_i, i = n_d + 1, \dots, n\}$. In most cases S_d consists of filters such as $\nabla_x, \nabla_y, \Delta$ which capture the general smoothness appearances of real world images, and S_r contains filters such as Gabor filters at various orientations and scales which characterize salient features of images.

Minimizing $U(\mathbf{I}; \Lambda, S)$ by gradient descent, we obtain the following non-linear parabolic partial differential equation:

$$\mathbf{I}_t = \sum_{i=1}^{n_d} F_i^- * \phi'_i(F_i * \mathbf{I}) + \sum_{i=n_d+1}^n F_i^- * \psi'_i(F_i * \mathbf{I}), \quad (7)$$

with $F_i^-(x, y) = -F_i(-x, -y)$. Thus starting from an input image $\mathbf{I}(x, y, t) = \mathbf{I}_0$, the first term diffuses the image by reducing the gradients while the second term forms patterns as the reaction forces as they favor large filter responses. So we called equation (7) the Gibbs Diffusion And Reaction Equation (GRADE).

Since the computation of equation (7) involves twice convolution for each of the selected filters, a conventional way for efficient computation is to build an image pyramid so that filters with large scales and low frequencies can be scaled down into small ones in the higher level of the image pyramid. This is appropriate especially when the filters are selected from a bank of multiple scales, such as the Gabor filters and the wavelet transforms.

For a given image \mathbf{I} , we build a pyramid with $\mathbf{I}^{(s)}$ being an image on a lattice $L^{(s)}$ at level $s = 0, 1, 2, \dots$ and $\mathbf{I}^{(0)} = \mathbf{I}$. Each pixel intensity of $\mathbf{I}^{(s+1)}$ is the average of a 2×2 pixels of $\mathbf{I}^{(s)}$. $F_i^{(s)}$ is a filter F_i used at level s .

Thus the potential function becomes,

$$U(\mathbf{I}; \Lambda, S) = \sum_s \sum_i \sum_{(x,y) \in L^{(s)}} \phi_i^{(s)}(F_i^{(s)} * \mathbf{I}_{(x,y)}^{(s)}) + \sum_s \sum_i \sum_{(x,y) \in L^{(s)}} \psi_i^{(s)}(F_i^{(s)} * \mathbf{I}_{(x,y)}^{(s)})$$

The reaction-diffusion equation can be easily derived for this pyramidal representation.

4 Properties of GRADE

This section studies some properties of the Gibbs reaction-diffusion equation.

First, we note that equation (7) can be considered as an extension to equation (1) on a discrete lattice by defining a vector field,

$$\vec{V}_{(x,y)} = (\phi'_1(), \dots, \phi'_{n_d}(), \psi'_{n_d+1}(), \dots, \psi'_n()),$$

and generalizing the divergence operator to

$$\text{div} = F_1^- * + F_2^- * + \dots + F_n^- *$$

Thus equation (7) can be written as

$$\mathbf{I}_t = \text{div}(\vec{V}). \quad (8)$$

Compared to equation (1) which transfers the “heat” among adjacent pixels, equation (8) transfers the “heat” in many directions on a graph, and the conductivities are defined as functions of the local patterns instead of just the local gradients. Note that in the discrete lattice, choosing $S_d = \{\nabla_x, \nabla_y\}$, $S_r = \emptyset$, we have $F_1 = F_1^- = \nabla_x$, $F_2 = F_2^- = \nabla_y$, and $\text{div} = \nabla_x + \nabla_y$, thus equation (1) and (2) are just special cases of equation (8).

Second, in figure 3, $\phi(\xi)$ has round tip for $\gamma \geq 1$, and a cusp occurs at $\xi = 0$ for $0 < \gamma < 1$ which leaves $\phi'(0)$ undefined, i.e. $\phi'(0)$ can be any value in $(-\infty, \infty)$ as shown by the dotted curves in figure 3d. An interesting fact is that the potential function learned from real world images does have cusp as shown in figure 2a where the best fit is $\gamma = 0.7$. This cusp forms because large part of objects in real world images have flat intensity appearances, and thus $\phi(\xi)$ with $\gamma < 1$ will produce piecewise constant regions with much stronger force than $\gamma \geq 1$.

By continuity, $\phi'(\xi)$ can be assigned any value in the range $[-\omega, \omega]$ for $\xi \in [-\epsilon, \epsilon]$ and $\omega = \phi'(\epsilon)$, as

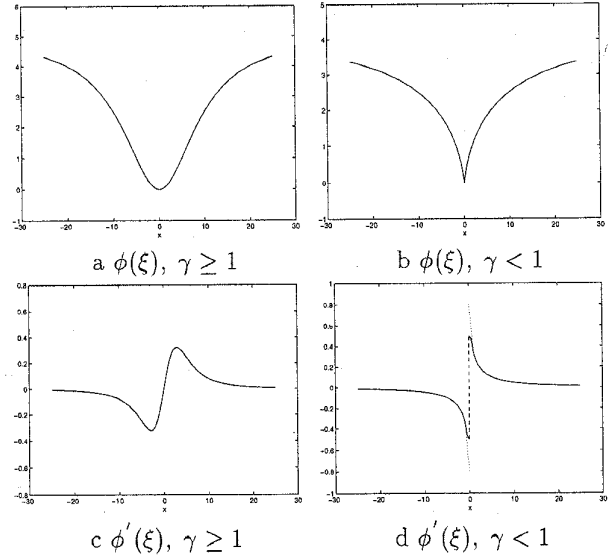


Figure 3: The potential function $\phi(\xi) = a(1 - \frac{1}{1 + (|\xi|/b)^\gamma})$ $a > 0$, and its derivative $\phi'(\xi)$ for a,c) $\gamma = 2.0$ and b,d) $\gamma = 0.8$.

shown by the dashed line in figure 3d. In numerical simulations, for $\xi \in [-\epsilon, \epsilon]$ we take

$$\phi'(\xi) = \begin{cases} +\omega & \text{if } \sigma < -\omega \\ -\sigma & \text{if } \sigma \in [-\omega, \omega] \\ -\omega & \text{if } \sigma > \omega \end{cases}$$

where σ is the summation of the other terms in the GRADE whose values are well defined. Intuitively when $\gamma < 1$ and $\xi = (F_i * \mathbf{I})(x, y) = 0$, $\phi'_i(0)$ forms an attractive basin in the neighborhood $\mathcal{N}_i(x, y)$ specified by the filter window F_i at (x, y) . For a pixel $(u, v) \in \mathcal{N}_i(x, y)$, the depth of the attractive basin is $\|\omega F_i^-(u - x, v - y)\|$. If a pixel is involved in multiple zero filter responses, it will accumulate the depth of the attractive basin generated by each filter. Thus unless the absolute value of the driving force from other well-defined terms is larger than the total depth σ of the attractive basin at (u, v) , $\mathbf{I}(u, v)$ will stay unchanged, i.e. $\mathbf{I}(u, v)$ will be locked at the current intensity by the attractive basin. In the image restoration experiments in later sections, $\gamma < 1$ shows better performance in forming piecewise constant regions.

Third, in contrast to existing diffusion equations, the GRADE includes reaction terms to produce patterns and to enhance preferred features. But unlike the Turing reaction-diffusion equation, the reaction terms in GRADE are defined intuitively by filters, and

their behavior is well specified by the learned Gibbs distribution. We will demonstrate this property in the experiments below.

5 Experiments

In this section, we demonstrate the applications of GRADE in pattern synthesis, denoising and image enhancement, and clutter removal.

5.1 Texture pattern formation

Two canonical patterns that the Turing reaction-diffusion equations can synthesize are the leopard blobs and zebra stripes [22][23]. In this section, we show that these patterns can be easily generated with only 2 or 3 filters using GRADE. We run equation (7) starting with $\mathbf{I}(x, y, 0)$ as a uniform noise image, and GRADE converges to a local minimum. Some synthesized texture patterns are displayed in figure 4.

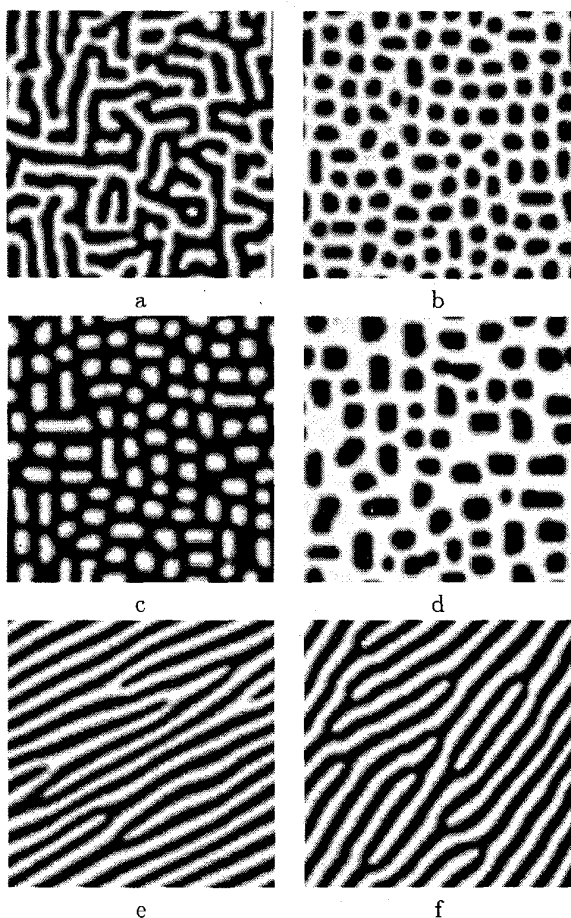


Figure 4: Leopard blobs and zebra stripes synthesized by GRADEs.

For all the six patterns in figure 4, we choose only

one diffusion filter $F_1^{(0)} = \Delta$ - the Laplacian of Gaussian filter at level 0 and we fix $a = 5, b = 10, \xi_o = 0, \gamma = 1.2$ for $\phi_1^{(0)}(\xi)$. For the three patterns in figure 4 a,b,c we choose isotropic center-surround filter $LG(\sqrt{2})$ of widow size 7×7 pixels as the reaction filter $F_2^{(1)}$ at level 1 of the image pyramid, and we set $(a = -6.0, b = 10, \gamma = 2.0)$ for $\psi_2^{(1)}(\xi)$. The differences between these three patterns are caused by ξ_o in $\psi_2^{(1)}(\xi)$. $\xi_o = 0$ forms the patterns with symmetric appearances for both black and white part as shown in figure 4a. As ξ_o becomes negative, black blobs begin to form as shown in figure 4b where $\xi_o = -6$, and positive ξ_o generates white blobs in the black background as shown in figure 4c where $\xi_o = 6$. The general smoothness appearance of the images is attributed to the diffusion filter. Figure 4d is generated with two reaction filters - $LG(\sqrt{2})$ at level 1 and level 2 respectively, thus blobs of mixed sizes appear. Similarly we selected one cosine Gabor filter $Gcos(4, 30^\circ)$ (size 7×7 pixels oriented at 30°) at level 1 as the reaction filter $F_1^{(2)}$ for figure 4e where $(a = -3.5, b = 10, \gamma = 2.0, \xi_o = 0)$ for $\psi_2^{(1)}(\xi)$. Figure 4f is generated with two reaction filters $Gcos(4, 30^\circ), Gcos(4, 60^\circ)$ at level 1 where $a = -3.5, b = 10, \gamma = 2.0, \xi_o = -3$ for $\psi_2^{(1)}(\xi)$ and $\psi_3^{(1)}(\xi)$.

It seems that the leopard blobs and zebra stripes are among the most canonical patterns which can be generated with easy choices of filters and parameters. As shown in [25], this new class of Gibbs distributions are capable of modeling a large variety of texture patterns, but specific forms for $\psi(\xi)$ have to be learned for a given texture pattern.

5.2 Denoising, edge enhancement, and clutter removal

So far we have studied the GRADE for minimizing a single energy function $U(\mathbf{I})$ by gradient descent. In image processing, we often need to model both the underlying images and the distortions, and to maximize a *posterior* distribution. Suppose the distortion is additive, i.e., an input image is,

$$\mathbf{I}^{\text{in}} = \mathbf{I} + \mathbf{C}.$$

In many applications, the distortion \mathbf{C} is often not i.i.d. Gaussian noise, but clutter with structures such as trees in front of a building or a military target. Such clutter will be very hard to handle by edge detection and image segmentation algorithms, and recently they have received increasing attention from people in various fields.

We propose to model clutter by an extra Gibbs distribution, which can be learned from some training

images by the minimax entropy theory as we did for the underlying image \mathbf{I} . Thus an extra pyramidal representation for $\mathbf{I}^{\text{in}} - \mathbf{I}$ is needed in a Gibbs distribution form as shown in figure 5. The filter set for the clutter is S_C . The resulting posterior distribution is still of the Gibbs form with potential function,

$$U^*(\mathbf{I}) = U_C(\mathbf{I}^{\text{in}} - \mathbf{I}; \Lambda_C, S_C) + U(\mathbf{I}; \Lambda, S), \quad (9)$$

where $U_C()$ is the potential of the clutter distribution.

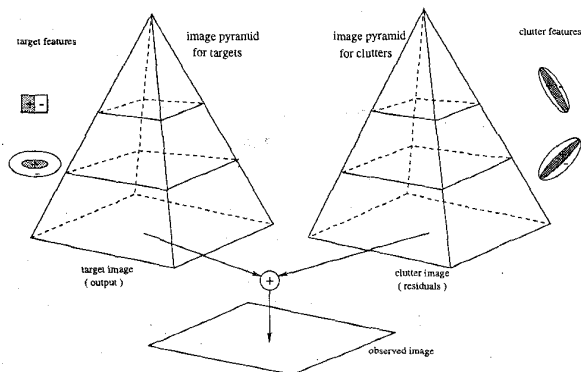


Figure 5: The computational scheme for removing noise and clutter.

Thus the MAP estimate of \mathbf{I} is the minimum of U^* . In the experiments which follow, we use the Langevin equation which adds perturbations to the gradient descent process and therefore is a variant of simulated annealing:

$$d\mathbf{I}_t = -\nabla U^*(\mathbf{I})dt + \sqrt{2T(t)}dw_t \quad (10)$$

where $w(x, y, t)$ is the standard Brownian motion process, i.e.,

$$w(x, y, t) \sim N(\mu(x, y), |t|), \quad dw_t = \sqrt{dt}N(0, 1).$$

$T(t)$ is the "temperature" which controls the magnitude of the random fluctuation. Under mild conditions on U^* , equation (10) approaches a global minimum of U^* at a low temperature. The analyses of convergence of the equations can be found in [9, 7, 5]. The computational load for the annealing process is notorious, but for applications like denoising, a fast decrease of temperature may not affect the final result very much.

Experiment I: Denoising and enhancing edge

In the first set of experiments, we take \mathbf{C} as an i.i.d. Gaussian noise image.

Figure 6a is image with additive i.i.d. Gaussian noise. We choose two diffusion filters $S_d = \{\nabla_x^{(0)}, \nabla_y^{(0)}\}$ at level 0, and fix $a = 5, b = 10, \gamma =$

$0.7, \xi_0 = 0$ for $\phi^{(0)}(\xi)$, ∇_x, ∇_y are also used as the reaction filters at level 1, 2, 3, i.e., $S_r = \{\nabla_x^{(1)}, \nabla_y^{(1)}, \nabla_x^{(2)}, \nabla_y^{(2)}, \nabla_x^{(3)}, \nabla_y^{(3)}\}$. We choose $a = -2.0, b = 10, \xi_0 = 0, \gamma = 1.6$ for $\psi_i^{(1)}()$, $a = -4.8, b = 15, \gamma = 2.0$ for $\psi_i^{(2)}()$, and $a = -10.0, b = 22, \gamma = 5.0$ for $\psi_i^{(3)}()$. Figure 6b is the recovered image.

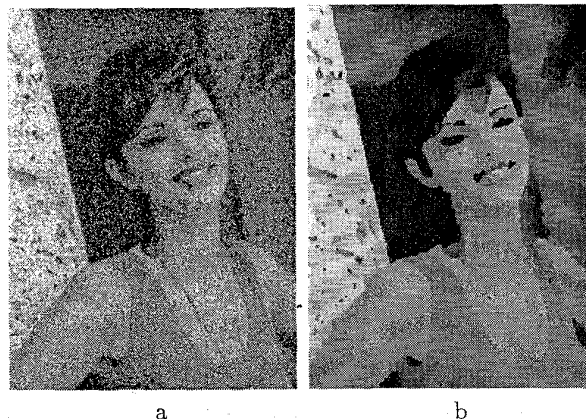


Figure 6: a. the image distorted with Gaussian noise, b. the restored image using the GRADE.

In figure 6b, noise is removed while boundaries are enhanced due to the reaction terms which have low energy for large gradients at the high levels of the pyramid, and this effect is not observed in the Perona-Malik anisotropic diffusion as in equation (2).

Experiment II: Clutter removal

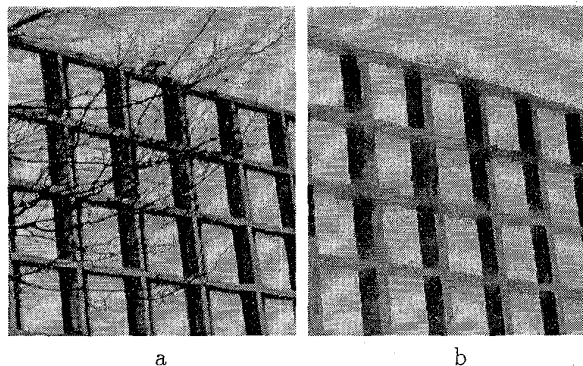


Figure 7: a. the observed image, b. the restored image using 6 filters.

Figure 7a shows tree branches in front of a building, which will make image segmentation and object recognition extremely difficult because they cause strong

edges across the image. Modeling such clutter is a challenging problem in many applications. In this paper, we only consider clutter as two dimensional pattern despite its geometry and 3D structures.

We collected a set of images of buildings and a set of images of trees all against clean background – the sky. For the tree images, we translate the image intensities to $[-31, 0]$, i.e., 0 for sky. In this case, since the trees are always darker than the buildings, then such negative intensity will approximately take care of the occlusion effects. We learn the Gibbs distributions for each set respectively in the pyramid, then such models are respectively adopted as the prior distribution and the likelihood as in equation (9). We recovered the underlying images by maximizing a posteriori distribution using the stochastic process.

For example, figure 7b is computed using 6 filters with 2 filters for \mathbf{I} : $\{\nabla_{x,0}, \nabla_{y,0}\}$, and 4 filters for \mathbf{I}_C : $\{\delta, \nabla_x, \nabla_y, G\cos(2, 30^\circ)\}$, i.e., the potential for \mathbf{I}_C is,

$$U_C(\mathbf{I}) = \sum_{(x,y)} \phi(\nabla_x \mathbf{I}(x,y)) + \phi(\nabla_y \mathbf{I}(x,y)) + \phi^*(\mathbf{I}(x,y)) + \psi^*(G\cos * \mathbf{I}(x,y)).$$

In the above equation, $\phi^*(\xi)$ and $\psi^*(\xi)$ are fit to the potential functions learned from the set of tree images,

$$\phi^*(\xi) = \begin{cases} a_1(1 - \frac{1}{1+(|\xi-\xi_0|/b)^\gamma}) & \xi < \xi_0 \\ a_2(1 - \frac{1}{1+(|\xi-\xi_0|/b)^\gamma}) & \xi \geq \xi_0, \end{cases} \quad a_2 >$$

So the energy term $\phi^*(\mathbf{I}(x,y))$ forces zero intensity for the clutter image while allowing for large negative intensities for the dark tree branches.

$$\psi^*(\xi) = \begin{cases} a(1 - \frac{1}{1+(|\xi-\xi_0|/b)^\gamma}) & \xi < \xi_0, \quad a > 0 \\ 0 & \end{cases}$$

Figure 8b is computed using 8 filters with 4 filters for \mathbf{I} and 4 filters for \mathbf{I}_C . 13 filters are used for figure 9 where the clutter is much heavier.

As a comparison, we run the anisotropic diffusion process [17] on figure 9a, and images at iterations $t = 50$ and ,300 are displayed in figure 10.

6 Discussion

Unlike previous PDEs motivated by chemical or physical processes, the GRADE is derived from Gibbs distribution, and the latter is learned from images in a given application. Therefore the GRADE is well tuned, and its behavior is well specified by the Gibbs distributions.

It is our belief that most of the PDEs currently used in image processing and computer vision can be unified under a general statistical framework, or can be

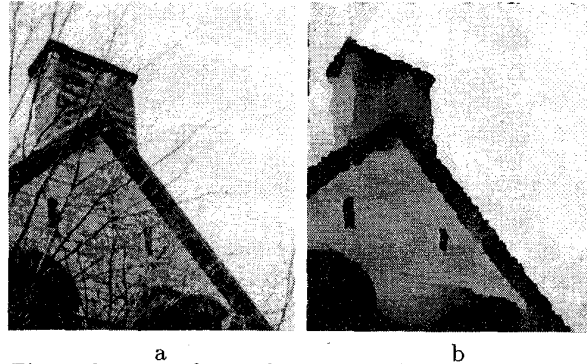


Figure 8: a. an observed image, b. the restored image using 8 filters.

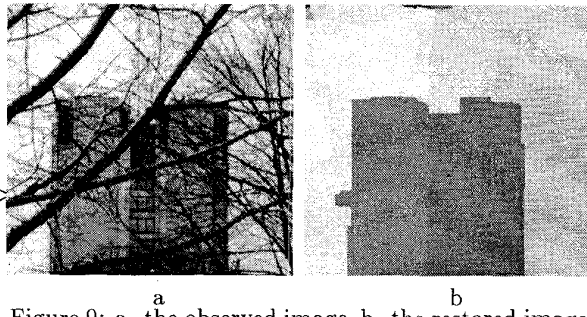


Figure 9: a. the observed image, b. the restored image using 13 filters.

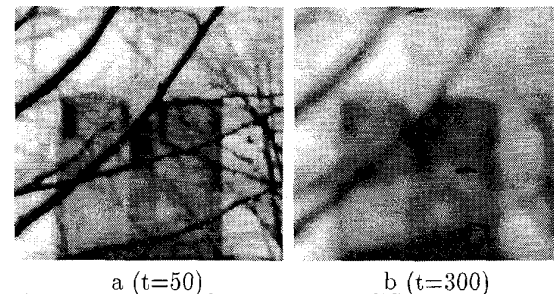


Figure 10: Images by anisotropic diffusion at iteration $t = 50, 300$.

replaced by PDEs of similar function derived from the statistical framework as we did. We note that the "region competition" algorithm studied in [24] is also a diffusion equation maximizing the posterior distribution for image segmentation. Similarly the scale space of shapes studied by Kimia etc. is also generated by a diffusion equation maximizing the prior distribution on a certain shape representation [11][15]. It is our hope that this work will stimulate further research efforts along this direction.

Acknowledgment:

This work is partially supported by an ARO grant DAAH04-96-1-0445 and an NSF grant to Mumford. For detailed information, contact szhu@cs.stanford.edu.

References

- [1] Black, M. Sapiro, G. Marimont, D and Heeger, D. "Robust anisotropic diffusion", *IEEE Trans. Image Processing*, (to appear).
- [2] Daugman, J. "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of Optical Society of America*. vol. 2, No. 7, pp1160-1169. 1985.
- [3] Gabor, D. "Theory of communication." *IEE Proc.* vol 93, no.26. 1946.
- [4] Geiger, D. and Giroso, F. "Parallel and deterministic algorithms for MRFs: surface reconstruction", *IEEE Trans PAMI*, 13(5):401-412, 1991.
- [5] Gelfand, S. B. and Mitter, S.K. "On sampling methods and annealing algorithms", in *Markov Random Fields - Theory and Applications* Academic Press, Inc, 1993.
- [6] Geman, S and Geman, D. "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images". *IEEE Trans. on PAMI* 6(7), pp 721-741, 1984.
- [7] Geman, S. and Hwang, C. "Diffusion for global optimization", *SIAM J. Control and Optimization*, Vol. 24, No. 5, September 1986.
- [8] Geman, S. and McClure, D. "Statistical methods for tomographic image reconstructions", *Bulletin of the International Statistical Institute*, 52, 4-20. 1987.
- [9] Gidas, B. "The Langevin equation as a global minimization algorithm" in *Disordered Systems and Biological Organizations*, Springer-Verlag, 1986.
- [10] Grindrod, P. *The theory and applications of reaction-diffusion equations*, Oxford Univ. Press, 1996.
- [11] Kimia, B., Tannebaum, A. and Zucker, S. "Shapes, shocks, and deformations I: the components of two-dimensional shape and the reaction-diffusion space", *Int'l J. Computer Vision*, vol. 15, pp189-224, 1995.
- [12] Mallat, S. "multi-resolution approximations and wavelet orthonormal bases of $L^2(R)$." *Trans. Amer. Math. Soc.* 315, 69-87. 1989.
- [13] Mumford, D. and Shah. J. "Optimal approximations by piecewise smooth functions and associated variational problems." *Comm. Pure Appl. Math.*, 42, pp 577-684, 1989.
- [14] Murray, J.D. "A pre-pattern formation mechanism for mammalian coat markings", *Journal of Theoretical Biology*, vol. 88, pp 161-199, 1981.
- [15] Niessen, W. etc. "A general framework for geometry-driven evolution equations", *Int'l J. computer Vision*, 21(3), 187-205, 1997.
- [16] Nitzberg, M. and Shiota, T. "Nonlinear image filtering with edge and corner enhancement" *IEEE. Trans. PAMI*, vol.14, Aug. 1992.
- [17] Perona, P and Malik, J, "Scale-space and edge detection using anisotropic diffusion" *IEEE Trans. on PAMI*, vol.12, No.7, July 1990.
- [18] Price, C.B., Wambacq, P. and Oosterlinck, A. "Applications of reaction-diffusion equations to image processing", *3rd Int'l Conf. on Image processing Applications*, July, 1989.
- [19] Sherstinsky, A and Picard, R. W. "M-lattice: from morphogenesis to image processing" *IEEE Trans. on Image Processing*, vol.5, no.7, July 1996.
- [20] Simoncelli, E. P. Freeman, W.T, Adelson, E. H, Heeger, D.J. "Shiftable multi-scale transforms. *IEEE Trans. on information theory*. Vol. 38, pp587-607, 1992.
- [21] Turing, A. M, "The chemical basis of morphogenesis", *Philosophy Trans. Royal Soc. London*, vol. 237, No. B, pp, 37-72, 1952.
- [22] Turk, G., "Generating textures on arbitrary surfaces using reaction-diffusion", *Computer Graphics*. Vol. 25, No.4, 1991.
- [23] Witkin, A. and Kass, M, "Reaction-diffusion textures", *Computer Graphics*. Vol. 25, No.4, 1991.
- [24] Zhu, S, and Yuille, A.L. "Region Competition: unifying snakes, r*IEEE Trans.on PAMI*. vol. 18, No.9, Sept. 1996.
- [25] Zhu, S, Wu, Y, and Mumford, D. B. "Filters, Random fields And Maximum Entropy (FRAME) - towards a unified theory for texture modeling", *Int'l J. of Computer Vision*, (to appear). A short version appeared in CVPR'96.
- [26] Zhu, S, and Mumford D.B. "Learning generic prior models for visual computation", *Proc. of Int'l Conf. on Computer Vision and Pattern Recognition*, 1997.