IEEE Trans. PAMI. vol.19, no.11, Nov. 1997. Short version in ICCV98.

GRADE: Gibbs Reaction And Diffusion Equation

– A framework for pattern synthesis, denoising, image enhancement, and clutter removal

Song Chun Zhu and David Mumford
Division of Applied Mathematics
Box F, Brown University
Providence, RI 02912
email zhu@dam.brown.edu

Abstract

This paper proposes a new class of nonlinear PDEs, called the *Gibbs Reaction And Diffusion Equation* (GRADE), for a variety of applications in computer vision, image processing, and graphics. In two previous papers, the authors have been studying a minimax entropy theory based on which a new class of Gibbs distributions are learned in a fully non-parametric form from a set of observed images such as textures and natural images. These Gibbs distributions have potentials of the form $U(\mathbf{I};\Lambda, S) = \sum_{\alpha=1}^{K}\sum_{(x,y)}\lambda^{(\alpha)}((F^{(\alpha)} * \mathbf{I})(x,y))$ with $S = \{F^{(1)}, F^{(2)}, ..., F^{(K)}\}$ being a set of filters and $\Lambda = \{\lambda^{(1)}(), \lambda^{(2)}(), ..., \lambda^{(K)}()\}$ the potential functions. We find that the learned $\lambda^{(\alpha)}()$s can be divided into two categories, and subsequently the partial differential equations given by gradient descent on $U(\mathbf{I};\Lambda, S)$ are essentially reaction-diffusion equations, where the energy terms in one category produce anisotropic diffusion while the energy terms in the second category produce reaction associated with pattern formation, enhancing preferred image features. This paper aims to provide a statistical framework of designing reaction-diffusion equations based on training images in a given application, instead of simulating physical and chemical processes as previous methods did. We demonstrate experiments where the GRADE is 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. In the past decade, 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 modeling the chemical mechanism for mammalian coat patterns (Turing 1952), such as leopard blobs and zebra stripes. Recently these PDEs are utilized to render realistic textures in graphics (Turk 1991, Witkin and Kass 1991), and they are also adopted in image processing such as enhancing fingerprints and halftoning (Price 1989, Sherstinsky and Picard 1996). (II) The Perona-Malik anisotropic diffusion equations (Perona and Malik 1990). These equations are used for generating scale space and removing noise (Nitsberg and Shiota 1992).

Despite the success of these PDEs, many questions are not understood in the literature. Firstly, the strengths of these PDEs lies on their nonlinearity, which also makes the equations unstable or unbounded especially when reaction terms are involved. Secondly these PDEs are derived from chemical or physical models, and there is little justification for applying them to images which seems irrelevant to the original chemical or physical processes. Thus the selection of the PDEs and their coefficients are subjective. Thirdly there lacks a generic theory to guide the design of PDEs for general-purposed applications in computer vision and image processing.

This paper proposes a statistical framework for designing a new class of PDE, called the *Gibbs Reaction And Diffusion Equation* (GRADE), inspired by our previous work on modeling texture and natural images. In two previous papers, the authors have been studying a minimax entropy theory based on which a Gibbs distribution $p(\mathbf{I})$ is learned from a set of observed images. $p(\mathbf{I})$ has potential of the form $U(\mathbf{I}; \Lambda, S) = \sum_{\alpha=1}^{K} \sum_{(x,y)} \lambda^{(\alpha)}((F^{(\alpha)} * \mathbf{I})(x, y))$ where $S = \{F^{(1)}, F^{(2)}, ..., F^{(K)}\}$ is a set of filters which best characterize the set of observed images, and $\Lambda = \{\lambda^{(1)}(), \lambda^{(2)}(), ..., \lambda^{(K)}()\}$ are the potential functions which are learned in a non-parametric form such that $p(\mathbf{I})$ reproduces the observed statistics captured by S. The learned Gibbs distributions were verified by stochastic sampling in the spirit of "analysis=synthesis", i.e., if $p(\mathbf{I})$ models the observed images correctly, the images sampled from $p(\mathbf{I})$ should have similar appearances to the observed ones.

We find that the learned $\lambda^{(\alpha)}()$s in $p(\mathbf{I})$ can be divided into two classes, and subsequently the partial differential equation given by gradient descent on $U(\mathbf{I};\Lambda, S)$ is essentially reaction-diffusion equation, which we named 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, enhancing preferred image features. Unlike the PDEs derived from a chemical or physical process, 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 analyzes its properties. Section (4) presents experiments on pattern synthesis, denoising, image enhancement, and clutter removal. Then section (5) concludes with a discussion.

2 Related work

This section reviews some of the nonlinear PDEs previously used in computer vision.

2.1 Reaction-diffusion model for texture pattern formation

A set of nonlinear PDEs was first studied in (Turing 1952) for modeling the formation of animal coat patterns by the diffusion and reaction of chemicals, which Turing called the "morphogens". These equations were further explored by Murray in theoretical biology (Murray 1981, 1993). 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 is by no means accepted by theoretical biologists because no such morphogen has been identified, however, they provide a way for synthesizing some texture patterns (Turk 1991, Witkin and Kass 1991). 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 reaction-diffusion equations are utilized in image processing such as enhancement of fingerprint images and image halftoning (Price 1989, Sherstinsky and Picard 1996).

Another set of PDEs for pattern formation was studied by Swindale in modeling the development of the binocular dominance stripes in the visual cortex of cats and monkey (Swindale 1980), and the simulated patterns are very similar to the zebra stripes. Again such theory is not verified in experiments of neurophysiology.

In the reaction-diffusion equations above, reaction terms are responsible for pattern formation, however they also make the equations unstable or unbounded, and human interference is needed. 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 rendering some patterns, and blobs and stripe appear to be the most canonic ones, 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 $\mathbf{I}(x, y, t)$ (Perona and Malik 1990). Their equation simulate the "heat" diffusion process,

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

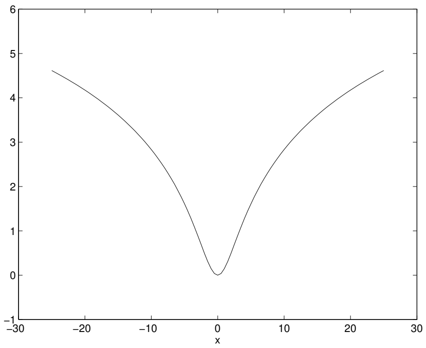
where $\nabla \mathbf{I} = (\mathbf{I}_x, \mathbf{I}_y)$ is the intensity gradient and div is the divergence operator, $\mathrm{div}(\vec{V}) = \nabla_x P + \nabla_y Q$ for any vector $\vec{V} = (P, Q)$. In practice, the heat conductivity $c(x, y, t)$ is defined as function of location gradients, for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

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

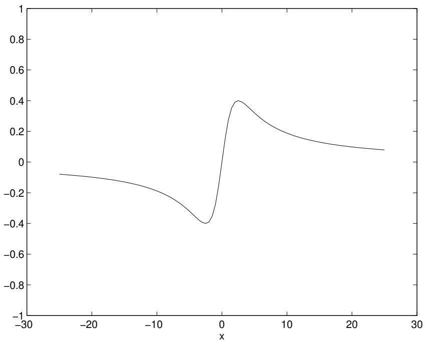
where b is a scaling constant. It is easy to see that equation (2) minimizes the following energy function by gradient descent,

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

where $\lambda(\xi) = const. \log(1+(\xi/b)^2)$ and $\lambda^{'}(\xi) = const. \frac{\xi}{1+(\xi/b)^2}$ are plotted in figure 1. Similar form of the energy functions are widely used as prior distributions in image restoration and segmentation (Geman and Geman 1984, Geman and McClure 1987, Blake and Zisserman 1987, Mumford and Shah 1989, Geiger and Girosi 1991).



a $\lambda(\xi)$



b $\lambda^{'}(\xi)$

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

Although the anisotropic diffusion equations can be adopted for removing noise and enhancing edges (Nitzberg and Shiota 1992), $\mathbf{I}(x, y, t)$ converges to a flat image as $t \to \infty$.

3 Gibbs reaction and diffusion equations

3.1 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 (Zhu, Wu and Mumford 1996, Zhu and Mumford 1996),

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

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

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

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 (Gabor 1946, Daugman 1985) and wavelet transforms (Mallat 1989, Simoncelli etc., 1992). 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 bank $\mathcal{B}$ of linear filters.

1. An intensity filter $\delta()$.

2. Isotropic center-surround filters, i.e., the Laplacian of Gaussian filters.

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

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, which are models for the frequency and orientation sensitive simple cells.

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

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 $Gcos(s, \theta)$ and $Gsin(s, \theta)$.

4. We will approximate large scale filters by filters of small window sizes on the high level of the image pyramid, where the image in one level is a "blown-down" version (i.e., averaged in 2×2 blocks) of the image blow.

In equation (3), the filter responses are quantitized into a finite number of bins and $\lambda_i(), i = 1, 2, ..., n$ becomes a piecewisely constant function or a vector. Thus the form of $\lambda_i(), i = 1, 2, ..., n$ 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 $F_i * \mathbf{I}^{obs}, i = 1, 2, ..., n$ over a set of training images $\mathbf{I}^{obs}$ in a given application.

The learned distribution is then verified by stochastic sampling, i.e., if $p(|I)$ models the given images correctly, the images sampled from $p(\mathbf{I})$ should have similar appearances to the training ones. A detailed account of this theory is referred to (Zhu, Wu and Mumford 1996).

In contrast to existing models on natural images and texture patterns in the literature, we found that the learned $\lambda_i(), i = 1, 2, ..., n$ can be divided into two classes. For example, two typical $\lambda_i()$s are plotted in figure 3.1, which are fit to a family of functions (see the dashed curves),

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

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 in contrast $\psi(\xi)$ has lower energy at the two tails which represent salient image features such as object boundaries.

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

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

Note that the filter set is divided into two parts $S = S_d \cup S_r$, with $S_d = \{F_i, i = 1, 2, .., n_d\}$ and $S_r = \{F_i, i = n_d + 1, ..., n\}$. 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

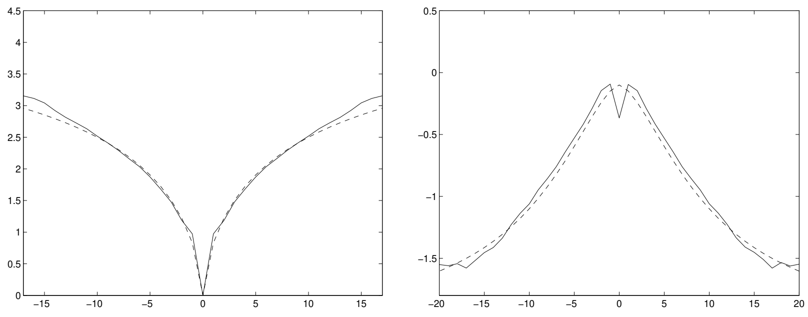


Figure 2 Learned potential functions $\lambda_i()$. The dashed curves are fitting functions: a). fitting $\phi(\xi)$ with $(a=5, b=10, \xi_o=0, \gamma=0.7)$. b,c,d). Fitting $\psi(\xi)$ with $(a=-2.0, b=10, \gamma=1.6)$.

---

partial differential equation:

$$\mathbf{I}_t = \sum_{i=1}^{n_d} F_i^- * \phi_i'(F_i * \mathbf{I}) + \sum_{i=n_d+1}^{n} F_i^- * \psi_i'(F_i * \mathbf{I}), \tag{7}$$

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

Since the computation of equation (7) involves convolving twice 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. We adopt this representation in our experiments.

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

3.2 Properties of the Gibbs diffusion-reaction equation

This section discusses 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,

$$\vec{V}_{(x,y)} = (\phi_1'(), ..., \phi_{n_d}'(), \psi_{n_d+1}'(), ..., \psi_n'()),$$

and generalizing the divergence operator to,

$$\mathrm{div} = F_1^- * + F_2^- * + \cdots + F_n^- * .$$

Thus equation (7) can be written as

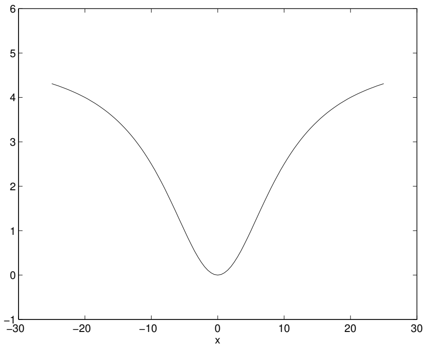
$$\mathbf{I}_t = \mathrm{div}(\vec{V}). \tag{8}$$

Compared to equation (1) which transfers the "heat" among adjacent pixels, equation (8) transfers the "heat" in many directions in 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 $F_1 = F_1^- = \nabla_x, F_2 = F_2^- = \nabla_y$, and $\mathrm{div} = F_1^- * + F_2^- *$, thus equation (1) and (2) can be derived from equation (8) by choosing $S_d = \{\nabla_x, \nabla_y\}, S_r = \emptyset$.

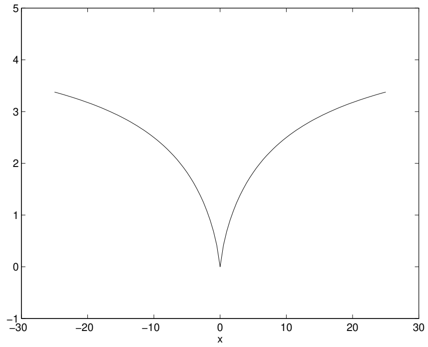
Second, in figure 3, $\phi(\xi)$ has round tip for $\gamma \geq 1$, and a cusp occurs at $\xi = 0$ for $0 < \gamma < 1$ which leaves $\phi'(0)$ undefined, i.e. $\phi'(0)$ can be any value in $(-\infty, \infty)$ as shown by the dotted curves in figure 3d. An interesting fact is that the potential function learned from real world images does have cusp as shown in figure 3.1a 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 $\omega = c\frac{\epsilon^{\gamma-1}}{(1+(\epsilon/b)^\gamma)^2}$, as shown by the dashed line in figure 3d. In numerical simulations, we take

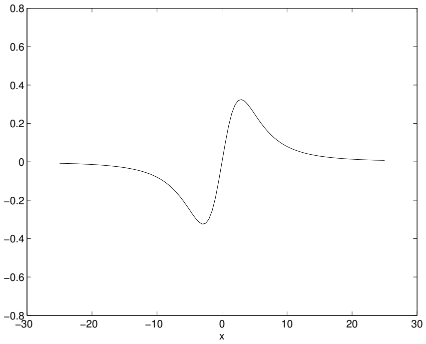
$$\phi'(0) = \begin{cases} +\omega & \text{if } \sigma < -\omega \\ -\sigma & \text{if } \sigma \in [-\omega, \omega] \\ -\omega & \text{if } \sigma > \omega \end{cases}$$



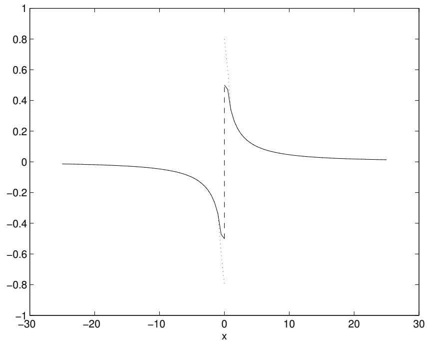
a $\phi(\xi),\ \gamma \geq 1$



b $\phi(\xi),\ \gamma < 1$



c $\phi^{'}(\xi),\ \gamma \geq 1$



d $\phi^{'}(\xi),\ \gamma < 1$

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) $\gamma = 0.8$.

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 of the attractive basin at (u, v), $\mathbf{I}(u, v)$ will stay unchanged. 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.

4 Experiments

In this section, we demonstrate the performances of the GRADE in experiments of pattern synthesis, denoising and image enhancement, and removing clutter.

4.1 Texture pattern formation

Two canonical patterns that the Turing reaction-diffusion equations can synthesize are the leopard blobs and zebra stripes (Turk 1991, Witkin and Kass 1991). In this section, we show that these patterns can be easily generated with only 2 or 3 filters using the 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.

For all the six patterns in figure 4, we choose the Laplacian of Gaussian filter Δ at level 0 of the image pyramid as the only diffusion filter $F_1^{(0)}$, 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

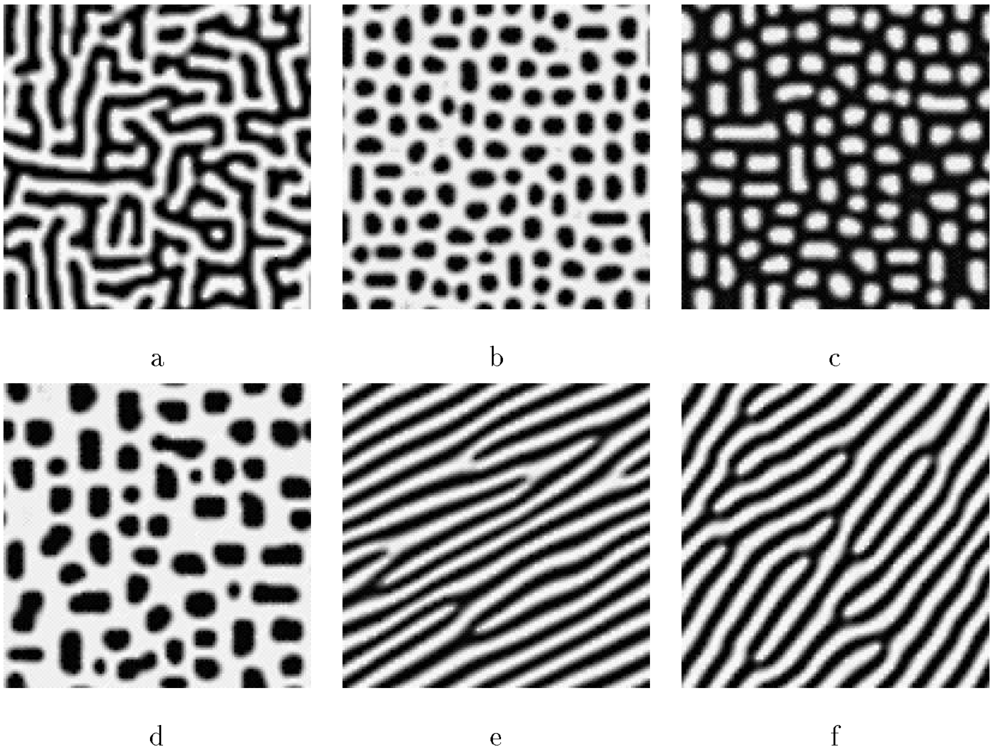


Figure 4 Leopard blobs and zebra stripes synthesized by GRADEs.

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 4b 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, therefore the GRADE creates blobs of mixed sizes. Similarly we selected one cosine Gabor filter $Gcos(4, 30^o)$ (size 7×7 pixels oriented at 30^o) 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 $Gcos(4, 30^o), Gcos(4, 60^o)$ 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 (Zhu, Wu, and Mumford 1996), 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.

4.2 Denoising, edge enhancement, 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 some 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 image $\mathbf{C}$ 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 resulting posterior distribution is still of the Gibbs form with potential function,

$$U^*(\mathbf{I}) = U_{\mathbf{C}}(\mathbf{I}^{\text{in}} - \mathbf{I}; \Lambda_{\mathbf{C}}, S_{\mathbf{C}}) + U(\mathbf{I}; \Lambda, S), \tag{9}$$

where $U_{\mathbf{C}}()$ is the potential of the clutter distribution.

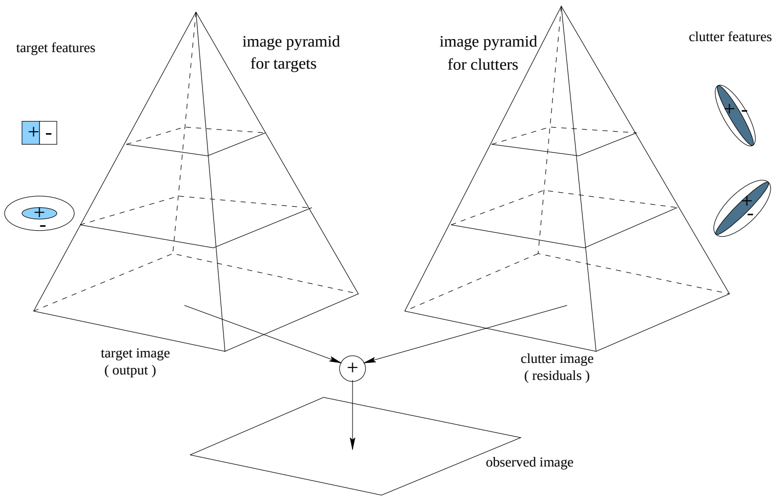


Figure 5 The computational scheme for removing noise and clutter.

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:

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

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

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

$T(t)$ 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 (Gidas 1986, Geman and Hwang 1986, Gelfand and Mitter 1993). 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

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

Figure 6a is an original image, and i.i.d. Gaussian noise is added in figure 6b. We choose two diffusion filters $S_d = \{\nabla_x^{(0)}, \nabla_y^{(0)}\}$ at level 0, and fix $a = 5, b = 10, \gamma = 0.7, \xi_o = 0$ 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 6d is the recovered image, as a comparison, figure 6c is the recovered image without the reaction filters, i.e., $S_r = \emptyset$.

In figure 6d, noise are removed while boundaries are enhanced due to the reaction terms which have low energy for large gradients at the high levels of the pyramid. We also run the Perona-Malik anisotropic diffusion as in equation (2), and figure 7 displays the images at iterations $t = 50, 100, 200$ starting with the noisy image in figure 6b. At the early iterations, the images removes the noise and object boundaries remains relatively sharp, but without the reaction terms and the term for noise, the boundaries will get blurred gradually.

Experiment II

In many applications, i.i.d. Gaussian models for distortions are not sufficient. For example, in figure 8a, the tree branches in the foreground 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

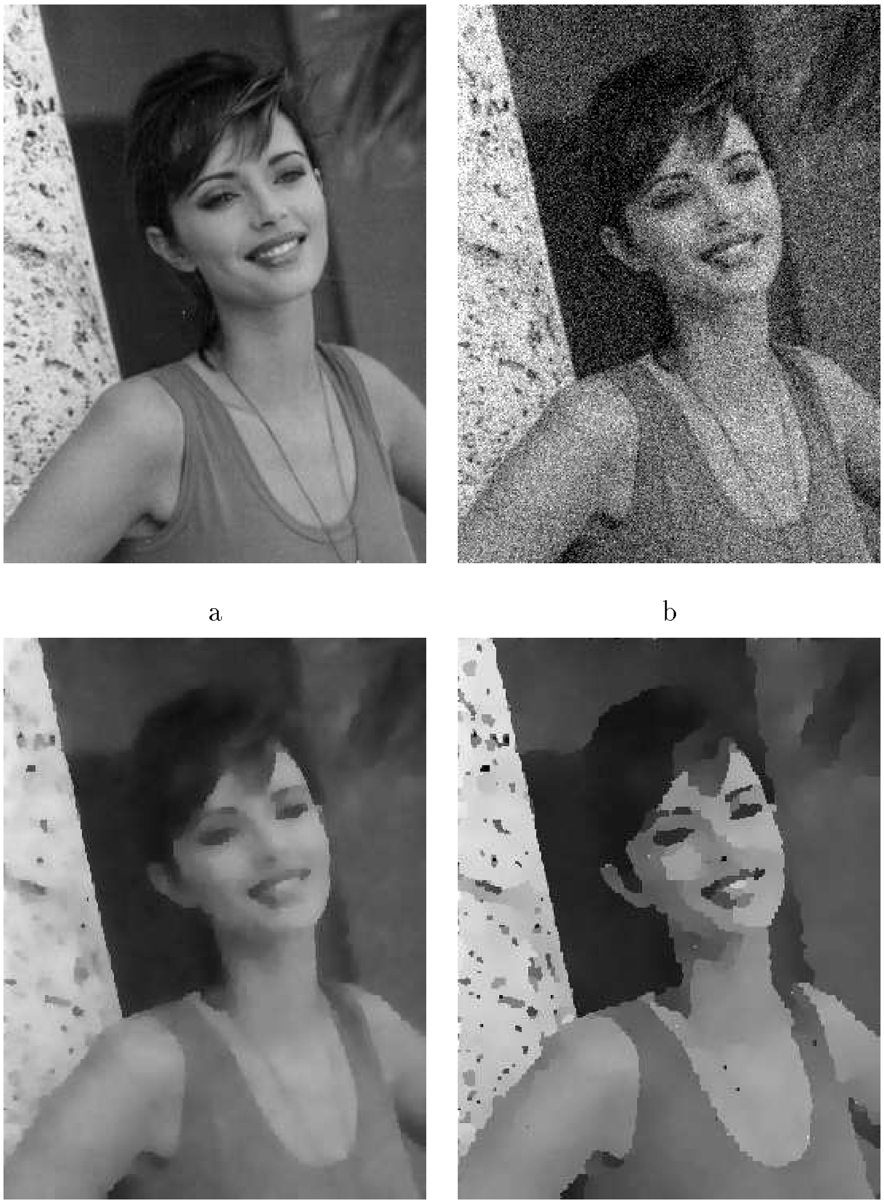
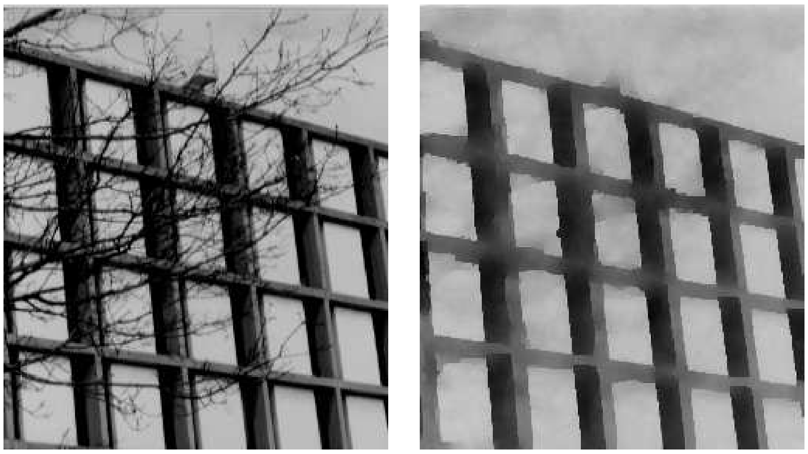


Figure 6 a. the original image, b. the image distorted with Gaussian noise, c. the restored image using diffusion terms only, d. the restored image using both diffusion and reaction terms.

a (t=50) b (t=100) c (t=200)

Figure 7 Images by anisotropic diffusion at iteration $t = 50, 100, 200$.



a b

Figure 8 a. the observed image, b, the restored image using 6 filters.

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 8b is computed using 6 filters with 2 filters for $\mathbf{I}$: $\{\nabla_{x,0}, \nabla_{y,0}\}$, and 4 filters for $\mathbf{I_C}$: δ, $\{\nabla_x, \nabla_y, Gcos(2, 30^o)\}$, i.e., the potential for $\mathbf{I_C}$ is,

$$U_{\mathbf{C}}(\mathbf{I}) = \sum_{(x,y)} \phi(\nabla_x \mathbf{I}(x,y)) + \phi(\nabla_y \mathbf{I}(x,y)) + \phi^*(\mathbf{I}(x,y)) + \psi^*(Gcos * \mathbf{I}(x,y)).$$

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

$$\phi^*(\xi) = \begin{cases} a_1(1 - \frac{1}{1+(|\xi-\xi_o|/b)^\gamma}) & \xi < \xi_o \\ a_2(1 - \frac{1}{1+(|\xi-\xi_o|/b)^\gamma}) & \xi \geq \xi_o, \quad a_2 > a_1 > 0. \end{cases}$$

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.

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

Figure 9b is computed using 8 filters with 4 filters for $\mathbf{I}$ and 4 filters for $\mathbf{I_C}$. 13 filters are used for figure 10 where the clutter is much heavier.

As a comparison, we run the anisotropic diffusion process (Perona and Malik 1990) on figure 10a, and images at iterations $t = 50, 100, 300$ are displayed in figure 11.

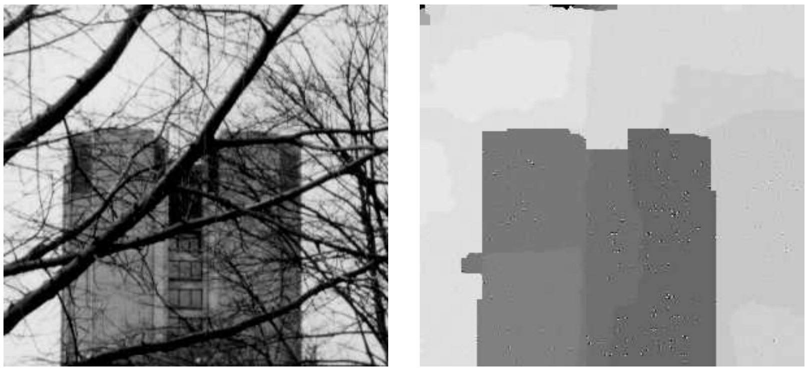
5 Discussion

This paper proposes a new class of nonlinear PDEs for texture pattern synthesis, denoising, image enhancement, and clutter removal. 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 replaced by PDEs of similar function derived from the statistical framework as we did. This paper aims to

a b

Figure 9 a. an observed image, b. the restored image using 8 filters.



a b

Figure 10 a. the observed image, b. the restored image using 13 filters.

a (t=50) b (t=100) c (t=300)

Figure 11 Images by anisotropic diffusion at iteration $t = 50, 100, 300$.

provide a generic method for designing PDEs at the first level, where no high level concept such as object boundary and geometrical properties are involved. In a proceeding paper, Zhu and Yuille studied an algorithm called "region competition", which is also a diffusion equation maximizing the posterior distribution for image segmentation (Zhu and Yuille 1996). 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 (Kimia, Tannebaum, and Zucker 1995). We call them the PDEs at second level vision. A survey of some diffusion equations are given in (Niessen etc 1997). It is our hope that this work will stimulate further research efforts along this direction.

References

[1] Blake, A. and Zisserman, A. *Visual Reconstruction.* MIT press, 1987.

[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 Girosi, F. "Parallel and deterministic algorithms for MRFs: surface reconstruction", *IEEE Trans PAMI,* 13(5):401-412, 1991.

[5] Geiger, D. and Yuille, A. "A common framework for image segmentation". *Inter.J.of Computer Vision* 6(3), pp 227-243 1991.

[6] Gelfand, S. B. and Mitter, S.K. "On sampling methods and annealing algorithms", in *Markov Random Fields – Theory and Applications* Academic Press, Inc, 1993.

[7] 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.

[8] Geman, S. and Hwang, C. "Diffusion for global optimization", *SIAM J. Control and Optimization*, Vol. 24, No. 5, September 1986.

[9] Geman, S. and McClure, D. "Statistical methods for tomographic image reconstructions", *Bulletin of the International Statistical Institute*, **52**, 4-20. 1987.

[10] Gidas, B. "The Langevin equation as a global minimization algorithm" in *Disordered Systems and Biological Organizations*, Springer-Verlag, 1986.

[11] Grindrod, P. *The theory and applications of reaction-diffusion equations*, Oxford Univ. Press, 1996.

[12] 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.

[13] Mallat, S. "multi-resolution approximations and wavelet orthonormal bases of $L^2(R)$." *Trans. Amer. Math. Soc.* **315**, 69-87. 1989.

[14] Mumford, D. and Shah. J. "Optimal approximations by piecewise smooth functions and associated variational problems." *Comm. Pure Appl. Math.,* **42**, pp 577-684, 1989.

[15] Murray, J.D. "A pre-pattern formation mechanism for mammalian coat markings", *Journal of Theoretical Biology*, vol. 88, pp 161-199, 1981.

[16] Murry, J. D. *Mathematical Biology*, Springer-Verlag, 1993.

[17] Niessen, W. etc. "A general framework for geometry-driven evolution equations", *Int'l J. computer Vision*, 21(3), 187-205,1997.

[18] Nitzberg, M. and Shiota, T. "Nonlinear image filtering with edge and corner enhancement" *IEEE. Trans. PAMI*, vol.14, Aug. 1992.

[19] Perona, P and Malik, J, "Scale-space and edge detection using anisotropic diffusion" *IEEE Trans. on PAMI*, vol.12, No.7, July 1990.

[20] 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.

[21] Sherstinsky, A and Picard, R. W. "M-lattice: from morphogenesis to image processing" *IEEE Trans. on Image Processing*, vol.5, no.7, July 1996.

[22] 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.

[23] Swindale, N.V. "A model for the formation of ocular dominance stripes", *Proc. R. Soc. Lond.* B 208, pp243-264, 1980.

[24] Turing, A. M, "The chemical basis of morphogenesis", *Philosophy Trans. Royal Soc. London*, vol. 237, No. B, pp, 37-72, 1952.

[25] Turk, G., "Generating textures on arbitrary surfaces using reaction-diffusion", *Computer Graphics.* Vol. 25, No.4, 1991.

[26] Witkin, A. and Kass, M, "Reaction-diffusion textures", *Computer Graphics.* Vol. 25, No.4, 1991.

[27] Zhu, S, and Yuille, A.L. "Region Competition: unifying snakes, region growing, and Bayes/MDL for multi-band image segmentation". *IEEE Trans.on PAMI.* vol. 18, No.9, Sept. 1996.

[28] 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.

[29] 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.