UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

On the Role of the BV Image Model in Image Restoration

Tony F. Chan
Jianhong Shen

April 2002

CAM Report 02-14

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90095-1555

http://www.math.ucla.edu/applied/cam/index.html

Contemporary Mathematics

On the Role of the BV Image Model in Image Restoration

Tony F. Chan and Jianhong Shen

ABSTRACT. What we believe images are determines how we take actions in image and lower-level vision analysis. In the Bayesian framework, it is manifest in the importance of a good image prior model. This paper intends to give a concise overview on the vision foundation, mathematical theory, computational algorithms, and various classical as well as unexpected new applications of the BV (bounded variation) image model, first introduced into image processing by Rudin, Osher, and Fatemi in 1992 [*Physica D*, 60:259-268].

1. Introduction: Image modeling

Image modeling, namely, finding a suitable way to describe and represent images, is perhaps the most fundamental and crucial step for the whole ladder of tasks in image and lower-level vision analysis. The underlying philosophy is, that the way we process and analyze images is often deeply influenced by what we believe (or *model*) they are.

As of February, 2002, the Google search engine returns about 50,000,000 documents containing the word "image." But this broad usage does not mean that we have already had a rigorous *mathematical* definition. In fact, even the Webster's Dictionary kicks the definition of "image" onto that of "picture," and then explains the latter vaguely as "a representation made by painting, drawing, or photography," which says nothing but only how "images" or "pictures" are formed. It reminds us of the concept of weight. Mankind had blindly used it for thousands of years until the giants Newton and Einstein first tried to decipher the meaning of gravity.

The mathematical challenge of image modeling roots in the diversity and complexity of images, from the rich geometric structures to a large dynamic range of scales. Most of us do not consider it a good idea to treat any function $u(x, y)$

1991 *Mathematics Subject Classification.* Primary 94A08; Secondary 68U10, 65K10.

Key words and phrases. Bounded variation, co-area, image representation, Bayesian, image model, data model, blur, noise, inpainting, error concealment, zooming, perceptual image coding.

Research supported by grants from NSF under grant number DMS-9973341 and from ONR under N00014-02-1-0015. Send all comments to Shen at the address shown in the end of the paper.

(equally) as an image. But it seems that no one has yet seized the right tool to characterize the boundary between images and non-images. Perhaps there is no such *sharp* boundary at all. That is expressed by the well known *Gibbs' Random Fields* model of Geman and Geman [**GG84**]. Based on filtering and statistical learning techniques, the model has been developed more generally by Zhu, Wu, and Mumford [**ZWM97, ZM97**]. Such stochastic approach for image modeling gets more theoretically mature in the very recent work of Mumford and Gidas [**MG01**] based on infinitely divisible law and axiomatization.

Apart from the stochastic theory of images, there is also the exploration of possible deterministic image models. Such transition is perhaps best described by Yves Meyer's very recent "$u + v$" notion [**Mey01**]. Here $v(x, y)$ represents the rapidly oscillatory component (noise or textures), or a stochastic sampling, while $u(x, y)$ captures the deterministic features.

In the very beginning of computer vision and artificial intelligence, Marr and his colleagues [**MH80**] already noticed the importance of *edges* for image understanding and visual communication. Edges are indeed an intrinsic feature for images since they define, segment, and correlate individual objects [**NMS93**]. Thus the deterministic component u should at least allow edges, or, one dimensional singularities, and cannot be a traditional Sobolev function. Mumford and Shah [**MS89**] singled out these edge features and proposed the famous object-edge free boundary image model. Recently Candes and Donoho [**CD**], and Pennec and Mallat [**PM00**] have been developing geometric wavelets such as *curvelets* and *bandlets* to model the component u (while leaving the oscillatory component v resonant with conventional wavelets).

Is there a simple linear functional space that legalizes edges and is easy to work with, but is not too loose to include too many "uninteresting" images. The answer was discovered by Rudin, Osher, and Fatemi [**Rud, ROF92, RO94**] in 1992. It is the Banach space of functions with bounded variations (BV), which allows jumps (or singularities with co-dimension 1) but also has a sufficient control over arbitrary oscillations. Ever since, the model has witnessed many applications in image denoising, deblurring, interpolation and inpainting, super-resolution and zooming, error concealment in wireless image transmission, medical imaging, and various inverse problems (see, for examples, [**ROF92, DS96, VO96, CW98, CS01d, COS01, CS01a**]). Also see Meyer's recent lectures [**Mey01**] detailing the role of BV in modeling the u-component in his $u + v$ notion.

The current paper attempts to give an overview on the theory and applications of Rudin, Osher, and Fatemi's BV image model for image restoration, with a special emphasis on our recent work of employing the BV image model as an image interpolant for the inpainting problem [**CS01a, CS01b, CS01c, CS02**].

Section II briefly lays out the Bayesian foundation for variational image restoration. Section III introduces the original Rudin-Osher-Fatemi TV restoration model. We present both its theory and computation. In section IV, we explain our recent effort on applying the BV image model as an efficient image interpolant for the inpainting problem, highlighted with several computational examples. The last section discusses two extra issues related to the BV image model and then concludes the paper.

2. Bayesian Framework for Image Restoration

If there indeed exists *the* most important principle in the entire field of computer and human vision analysis, it has to be the Bayesian rule.

Many problems in image and vision analysis can be set up as follows. We are to infer some feature or patten (vector or continuous field) F from a given measured or observed data field X_0. For example, for image restoration, X_0 corresponds to a given corrupted image u_0, which is often blended with noise, blurred by de-focusing or medium scattering, or has certain data missing during the transmission process; and F denotes the ideal image u that one would get without all those distortion effects. For vision analysis, X_0 may represent the 2-D image u, while F denotes the 3-D configuration parameters (illuminance and reflectance, etc.) [**Ker**].

The ideal inference of F from X_0 is naturally the one that maximizes the *posterior* probability $\text{Prob}(F|X_0)$. According to the Bayes formula

$$\text{Prob}(F|X_0) = \frac{\text{Prob}(X_0|F)\ \text{Prob}(F)}{\text{Prob}(X_0)},$$

it suffices to maximize the product of the data model $\text{Prob}(X_0|F)$ and the prior model $\text{Prob}(F)$, since the denominator is merely a normalization constant once X_0 is given. The prior model $\text{Prob}(F)$ specifies how often a pattern F can be observed *a priori* , i.e., independent of any observation made. The data model $\text{Prob}(X_0|F)$ then reveals the likelihood for X_0 being generated from a given pattern F.

If one has the *a priori* evidence for the importance of *geometric* structures in the pattern distribution $\text{Prob}(F)$ (such as edges and their geometry for image understanding), then it is more convenient to work with the "energy" form of the Bayesian method, as Mumford did for various segmentation models [**Mum94**]. This is at least formally achieved via Gibbs' formula in statistical mechanics [**Gib02**]: the likelihood for a configuration F being observed is associated to its energy $E[F]$ by

$$\text{Prob}(F) = \frac{1}{Z} \exp\left(-\frac{E[F]}{kT}\right),$$

where k and T denote the Boltzmann constant and absolute temperature, and Z the partition function over all the permissible configurations. The meaning of the

energy $E[X_0|F]$ is similarly defined, though lacking a rigorous counterpart in statistical mechanics. Therefore, the Bayesian method leads to the energy minimization problem

$$\min_F E[F] + E[X_0|F].$$

In the literature of deterministic inverse problems, this corresponds to the celebrated idea of Tikhonov regularization [**Tik63**]. The Bayesian approach is more general in many aspects.

In terms of image restoration $u_0 \to u$, we are to minimize

$$E[u] + E[u_0|u]. \tag{1}$$

The data model $u \to u_0$ depends on the real physical imaging process. One popular and useful choice as in astronomic and many medical imaging processes is blurring followed by noising [**ROF92, CW98**]:

$$u_0 = Ku + n,$$

where n denotes additive noise, and the linear operator K models the blurring process

$$Ku(x) = \int_\Omega K(x,y)u(y)dy.$$

K is *lowpass* in the sense that $K1 \equiv 1$. As is well known in signal processing, if K is *shift-invariant* (or spatially homogeneous), then it has to be an ordinary filter $h(x)*$ via the convolution operator: $K(x,y) = h(x-y)$. Although realistic blurring factors fluctuate randomly, most often we observe that K is fixed deterministically. We shall do so in this paper as well. Modeling the white noise by Gaussian, we easily obtain the energy for the data model (up to a multiplier):

$$E[u_0|u] = \int_\Omega \frac{(Ku(x) - u_0(x))^2}{\sigma^2(x)}dx, \tag{2}$$

where $\sigma^2(x)$ is the noise variance at pixel x, and is constant for homogeneous noise. Generally $1/\sigma^2(x)$ simply contributes as a positive weight $w(x)$, and the energy presents a *weighted* least square fitting as discussed in [**GL83, Str93**].

Other blurring and noising models are also possible depending on the real imaging processes. For example, Rudin and Osher also studied the multiplicative noise model in [**RO94**].

Therefore, the restoration quality by (1) crucially depends on the choice of the image model $E[u]$. The BV image model of Rudin-Osher-Fatemi [**ROF92**] captures the edge feature of images, and is perhaps the most efficient geometric image model in terms of theoretical accessibility, computational efficiency, and applicational quality.

3. The BV Model of Rudin, Osher, and Fatemi

3.1. Functions with bounded variation. We start with some essential mathematical theory on functions with bounded variations. We refer to the outstanding monograph by Giusti [**Giu84**] for more details.

Let $\Omega \subset R^2$ denote the open image domain, which for most real applications bears a rectangular shape. For each real function $f \in L^1_{\text{loc}}(\Omega)$ (i.e., locally integrable), its total variation $\text{TV}(f)$ is defined in the distributional sense:

$$\text{TV}(f) = \sup_{\vec{g} \in C_0^1(\Omega, B_2)} \int_\Omega f(\nabla \cdot \vec{g})\, dx, \tag{3}$$

where B_2 denotes the unit disk in R^2, and the space of test functions is

$$C_0^1(\Omega, B_2) = \{\text{all } C^1 \text{ maps } \vec{g} : \Omega \to B_2, \text{ which are compactly supported}\}.$$

Since $f \in L^1_{\text{loc}}(\Omega)$ and $\nabla \cdot \vec{g} \in C_0(\Omega)$, $\text{TV}(f)$ is well defined. $\text{TV}(f) \geq 0$ since B_2 is closed under reflection $x = (x_1, x_2) \to -x$. Suppose f is in the Sobolev space $W^{1,1}(\Omega)$, then $\nabla f \in L^1(\Omega)$ and

$$-\int_\Omega f \nabla \cdot \vec{g}\, dx = \int_\Omega (\nabla f) \cdot \vec{g}\, dx,$$

which immediately implies that

$$\text{TV}(f) = \int_\Omega |\nabla f| dx = \int_\Omega \sqrt{f_{x_1}^2 + f_{x_2}^2}\, dx_1 dx_2.$$

It is for this reason that $\text{TV}(f)$ is often denoted by $\int_\Omega |Df|$, with the symbol D referring to the conventional differentiation ∇, and the absence of the Lebesgue area element dx indicating that $|Df|$ is a general Radon measure.

The Rudin-Osher-Fatemi image model takes $E[u] = const. \times \text{TV}(u)$ [**RO94, ROF92**], and applies to the class of images with bounded variations. The space of functions with *bounded variations* is defined as

$$\text{BV}(\Omega) = \{f : f \in L^1(\Omega) \text{ and } \text{TV}(f) < \infty\}.$$

It can be easily shown that $\text{BV}(\Omega)$ is a Banach space under the BV norm

$$\|f\|_{\text{BV}} = \|f\|_{L^1} + \text{TV}(f),$$

and it is continuously embedded in $L^1(\Omega)$.

Among all the important properties, there are three that have helped Rudin-Osher-Fatemi's BV image model become easily accessible in theory, and meaningful for applications in image and low-level vision analysis: (a) lower semi-continuity, (b) weak L^1 compactness, and (c) the co-area formula.

Lower semi-continuity of the TV norm in $L^1(\Omega)$ says if $f_n \to f$ weakly in $L^1(\Omega)$, then

$$\int_\Omega |Df| \leq \liminf_{n \to \infty} \int_\Omega |Df_n|.$$

In addition, the embedding $\mathrm{BV}(\Omega) \to L^1(\Omega)$ is compact, i.e., the unit ball of $\mathrm{BV}(\Omega)$ is compact in $L^1(\Omega)$. As well practiced in the direct method in Calculus of Variations, these two key properties together often point to the existence of minimizers for energies involving the TV norm. This is indeed the case in Chambolle and Lions' work on the TV restoration model [**CL97**], which will be outlined in the next section.

The third property reveals the geometric nature of the TV norm, and thus strongly supports its application in geometry motivated vision and image analysis. It is the *co-area* formula. Define the perimeter $\mathrm{Per}(Q)$ of a domain $Q \subset \Omega$ to be

$$\mathrm{Per}(Q) = \int_\Omega |D1_Q(x)| = \mathrm{TV}(1_Q),$$

which generalizes the conventional notion of length for a regular boundary ∂Q. For any function $u \in \mathrm{BV}(\Omega)$, the co-area formula [**Giu84**] says

$$\int_\Omega |Du| = \int_{-\infty}^{\infty} \mathrm{Per}(u < \lambda)\, d\lambda. \tag{4}$$

Here the event $(u < \lambda)$ denotes the domain $Q_\lambda = \{x \in \Omega : u(x) < \lambda\}$.

To better understand the formula, imagine a simple case when $u \in C^\infty$ and λ is regular, i.e., $\nabla u(x)$ does not vanish on the entire level set $u = \lambda$. Then the boundary ∂Q_λ is a regular smooth curve and $\mathrm{Per}(Q_\lambda)$ is exactly its Euclidean length. Therefore, in the conventional sense, the co-area formula states that

$$\boxed{\mathrm{TV}(u) \text{ is a collective way to sum up the lengths of all level lines.}}$$

It is this property that brings the TV norm closer to meeting the requirement of an ideal vision measure. Generally, human vision tends to represent curves and edges as simple as possible for the purpose of efficient neuronal data compression and visual communication [**Don00**]. Such representation is achieved by having the local small ripples ignored or filtered out, just as having the curve lengths shortened. This is the vision rationale for the minimization of the TV norm and the BV image model.

3.2. TV restoration: Model and theory. Section 2 and 3.1 lay out the vision and mathematical foundations for the original restoration model of Rudin, Osher, and Fatemi [**ROF92, RO94**].

As in Section 2, assume that a given image u_0 is noisy and blurred:

$$u_0 = Ku + n,$$

and in addition, the ideal image u is assumed in $\mathrm{BV}(\Omega)$. Then the Bayesian restoration energy first proposed by Rudin, Osher, and Fatemi [**ROF92, RO94**] is

$$E[u|u_0] = \alpha \mathrm{TV}(u) + E[u_0|u],$$

where α is a constant and the data model $E[u_0|u]$ is given as in (2). More explicitly and cleanly, up to a constant multiplier, we are to minimize

$$E[u|u_0] = \int_\Omega |Du| + \int_\Omega (Ku - u_0)^2 w(x)dx, \tag{5}$$

with $w(x) = 1/[\alpha\sigma^2(x)]$. In the original paper [**ROF92**], the noise is assumed to be homogeneous, and thus $w(x) = \lambda$ is a constant weight, with λ taking the effect of a Lagrange multiplier.

The existence and uniqueness of the TV restoration model in $\mathrm{BV}(\Omega) \bigcap L^2(\Omega)$ were proven by Chambolle and Lions using the direct method [**CL97**]. The major properties leading to the proof are the *lower semi-continuity* and L^1 *compactness* as outlined in the previous section. The basic assumptions ensuring the existence and uniqueness are

(a) (blurring model) The linear blurring operator $K : L^2(\Omega) \to L^2(\Omega)$ is continuous, *lowpass*: $K1 \equiv 1$, and *injective* (for uniqueness).
(b) (noise model) The white noise n has mean 0 and variance σ^2, known *a priori*.
(c) (independence of blurring and noise) $\mathrm{Var}(u_0) \geq \sigma^2$.

We should say a few more words about the last condition. In the current data model, we have assumed that the blurring K and noise n are independent. Therefore, from probability,

$$\mathrm{Var}(u_0) = \mathrm{Var}(Ku) + \mathrm{Var}(n) = \sigma^2 + \mathrm{Var}(Ku) \geq \sigma^2.$$

(If both the blurring K and the ideal image u are deterministic, then equality is indeed achieved.) But from the application point of view, most often we are only given one single observation u_0, despite that u_0 is a random field. Therefore, the last condition is numerically understood and inspected in the ergodic sense:

$$\mathrm{Var}(u_0) = \frac{1}{|\Omega|}\int_\Omega \left(u_0 - \frac{1}{|\Omega|}\int_\Omega u_0 dx\right)^2 dx,$$

where $|\Omega|$ denotes the Lebesgue measure of the image domain.

Recently, the TV restoration model (5) has been extended to data that live on general graphs (the so-called *digital TV*) [**COS01**], and to "non-flat" data or image features (such as chromaticity, orientations of optical flows, and orthogonal frames) that live on Riemannian manifolds [**CS01d, TSC, TD01**]. Strong and Chan also have studied the scale dependency and edge preserving properties of the TV model in much more details [**SC**].

3.3. TV restoration: Computation and approximation. To computationally realize the TV restoration model (5), as first proposed by Rudin, Osher, and Fatemi [**RO94**], one typically takes the steepest descent method (time marching; also see the recent work [**MO00**] for example) or directly solves the associated equilibrium equation (steady solution) by iterative methods [**VO96, DV97**]. Here we

discuss the latter. Other possible techniques include, for example, the primal-dual method of Chan, Golub, and Mulet [**CGM99**].

Formally, or assuming u in a finer space $H^1(\Omega)$, we find that the equilibrium state of energy (5) satisfies the Euler-Lagrange equation

$$\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) - 2K^* w(Ku - u_0) = 0, \tag{6}$$

with the Neumann adiabatic boundary condition. Here K^* is the adjoint of K, and $w(x) \equiv \lambda$ corresponds to the homogeneity of the noise in the original Rudin-Osher-Fatemi model. If indeed $u \in H^1(\Omega)$, then the differential equation is understood in the weak sense as in the classical theory of elliptic equations (i.e., in $(H^1)'$).

For a smooth function u, the differential term $\nabla \cdot [\nabla u/|\nabla u|]$ in (6) at a regular (i.e. with non-degenerate gradient) pixel x_0 is exactly the curvature of the level line $u \equiv u(x_0)$, which once more reveals the geometry encoded into the model.

On regions where the image is very smooth or near constant, the denominator $|\nabla u|$ blows up the diffusivity coefficient in (6). Therefore, most of the existing numerical algorithms have depended on the relaxation of the TV norm. That is,

$$\text{to replace} \quad \int_\Omega |Du| \quad \text{by} \quad \int_\Omega \phi_\epsilon(|\nabla u|)\, dx.$$

Here $\phi_\epsilon(p) : R \to R$, with $\phi'_\epsilon(p) > 0$, for $p > 0$ is a C^1 convex even function that mollifies that original $|p|$, and ϵ is a small relaxation parameter that often models the sensitivity of a vision system. Popular choices include

$$\phi_\epsilon(p) = \sqrt{p^2 + \epsilon^2} := |p|_\epsilon,$$

and the integral ϕ_ϵ of $f_\epsilon(p) = p\,(\epsilon \vee p^{-1} \wedge \epsilon^{-1})$ with $\phi_\epsilon(0) = 0$. (Here the wedges represent the ceiling and flooring operators [**CL97**]) Consequently, the equilibrium equation for such a relaxation is modified to

$$\nabla \cdot \left(\frac{\phi'_\epsilon(|\nabla u|)}{|\nabla u|} \nabla u \right) - K^* w(Ku - u_0) = 0. \tag{7}$$

It can be shown that the solutions to these two relaxed problems are all in $H^1(\Omega)$.

Computationally, the nonlinear equation (6) is often solved iteratively by the freezing-coefficient technique. That is, at each step n, the next update $u^{(n+1)}$ solves the linearized Poisson equation with blurring and fitting:

$$\nabla \cdot \left(z^{(n)} \nabla u \right) - K^* w(Ku - u_0) = 0, \tag{8}$$

where $z = \phi'_\epsilon(|\nabla u|)/|\nabla u| > 0$ is the diffusivity coefficient, and is freezed at the current step (see Acar and Vogel [**AV94**]). Therefore, from the energy point of view, the update is the unique minimizer to the elliptic energy

$$\int_\Omega z^{(n)} |\nabla u|^2 dx + \int_\Omega (Ku - u^0)^2 w(x) dx$$

in $H^1(\Omega)$. The convergence of such algorithms has been confirmed in [**AV94, CL97, CM, DV97**]. Furthermore, it is even possible to endow an energy meaning to the updating of z itself: $z^{(n+1)} = \phi'_\epsilon(|\nabla u^{(n+1)}|)/|\nabla u^{(n+1)}|$. For instance, for the choice of $\phi_\epsilon(p) = |p|_\epsilon$, it is easy to show that $z^{(n+1)}$ is the minimizer of

$$\int_\Omega \left[z\,|\nabla u^{(n+1)}|_\epsilon^2 + \frac{1}{z} \right] dx.$$

(For the other example, see Chambolle and Lions [**CL97**].) Therefore, in this case, the iterative algorithm based on the freezing-coefficient technique is essentially to minimize

$$E_\epsilon[u, z|u_0] = \int_\Omega \left[z(|\nabla u|^2 + \epsilon^2) + \frac{1}{z} \right] dx + \int_\Omega (Ku - u_0)^2\, w(x) dx.$$

Unlike the original Rudin-Osher-Fatemi model [**ROF92, RO94**], it now contains a new feature z, which is often called the *auxiliary variable* in the vision community [**GR92**]. We also call z the *edge signature* since if z is plotted as an image, then the dark (i.e., small z) and thin (depending on ϵ) stripes clearly outline the edges in the image.

4. BV as an Interpolant: Image Inpainting

4.1. The problem of inpainting. The word "inpainting" is an artistic synonym for "image interpolation," as initially circulated among museum restoration artists, who manually recover the cracks of degraded ancient paintings by following as faithfully as possible the intention of their original creators. The concept of "digital inpainting" was first introduced into image processing in the paper by Bertalmio, Saprio, Caselles, and Ballester [**BSCB00**], where a third order transport type nonlinear PDE was invented for the inpainting problem.

Recently, the concept of inpainting has been connected to many major problems in image processing and low-level vision, such as perceptual image coding and compression, and error concealment for wireless image transmission [**CS01a**]. We refer to our recent survey paper [**CS02**] for much more details on the status of the inpainting problem.

Traditionally, image interpolation is often restricted to problems with scattered *small-scale* missing data. Thus the approaches and algorithms have been mostly developed from the viewpoints of the spectral method, filtering method, wavelets, and radially symmetric bases, etc [**AG01, BTU01, IP97, LO01**]. But for *large-scale* interpolation, or "inpainting," these conventional approaches do not seem to work well due to the fundamental challenge: how to faithfully (at least visually meaningfully) recover the missing edges, i.e., the 1-dimensional singular feature of images.

In a different direction, Bertalmio et al. [**BSCB00**] have recently introduced the idea of applying transport type high order PDEs to complete the broken level lines. The authors of the present paper then took a different approach by having the inpainting problem embedded into the general category of image restoration. As a result, our inpainting models have been based on the Bayesian framework of Section 2. The first image model catching our attention was the BV image model of Rudin, Osher, and Fatemi. This is the TV inpainting model that we first proposed and studied in [**CS01a**].

As one shall see in the next section, TV inpainting is almost identical to the original TV restoration model (5). The beauty lies in that the slight modification dramatically extends its scope of applicability, and reveals many unexpected connections to other important problems in image and low-level vision analysis, such as perceptual image coding and super-resolution [**CS01a**].

After the TV inpainting was first introduced, our further recent works have demonstrated that as a special restoration problem, inpainting does carry its own identity and important differences from the more familiar types such as denoising and deblurring. In [**CS01c, CKS01, ES02**], we have shown that the BV image model is insufficient for large scale inpainting problems, and high order geometric image models based on curvatures are necessary for more faithful reconstruction of partially missing edges. The key reason for this is that TV only involves the first order geometric information (i.e. the length) of level-lines, but curvature also plays a crucial role in human visual interpolation [**Kan79**].

Nevertheless, the BV image model still remains one of the simplest and most effective image interpolants, and TV inpainting is among the very few inpainting models that are readily open to both theoretical analysis and efficient computational implementation. And even for realistic applications, it always provides a valuable lower order initial guess for computationally more expensive high order models. We now explain TV inpainting and some of its major applications in digital image processing.

4.2. The TV inpainting model by Chan and Shen. Let $D \subset \Omega$ denote the compact inpainting domain, on which the observation $u_0\big|_D$ is missing. The goal of inpainting is to recover the ideal image u on the *entire* domain Ω based on the available portion $u_0\big|_{\Omega\setminus D}$.

It is quite obvious that generally the inpainting problem is ill-posed: without the input from high-level vision operators, such as symmetry detection or more general pattern learning, it is impossible to inpaint an object that is completely missing. However, the stroke of luck does shine in many major digital applications [**CS01a**]

(e.g., random packet loss in wireless image transmission, image zooming and super-resolution, etc.) in that $u_0\big|_{\Omega\setminus D}$ indeed retains crucial information about $u_0\big|_D$.

The generative data model for inpainting is

$$u_0\big|_D = (Ku + n)\big|_D.$$

Under the assumptions in Section 2, it leads to the energy form

$$E_D[u_0|u] = \int_{\Omega\setminus D} (Ku - u_0)^2 w(x)dx,$$

with the weight $w(x) = const./\sigma^2(x)$. Let $w^D(x)$ denote the zero extension of $w(x)\big|_{\Omega\setminus D}$ onto the whole domain Ω, i.e.,

$$w^D(x) = (1 - 1_D(x))w(x).$$

Then the energy for the data model can also be written as

$$E_D[u_0|u] = \int_{\Omega} (Ku - u_0)^2 w^D(x)dx, \tag{9}$$

where $u_0\big|_{\Omega\setminus D}$ is extended to $u_0 = u_0\big|_{\Omega}$ in any manner since it is wiped out by $w^D(x)$ anyway.

Then Bayesian inpainting based on the BV image model is to minimize

$$E_D[u|u_0] = \int_{\Omega} |Du| + \int_{\Omega} (Ku - u_0)^2 w^D(x)\, dx, \tag{10}$$

which is almost identical to the original TV restoration model, only with an adjustment on the weight function. Consequently, they share the same form of the Euler-Lagrange PDE:

$$\nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right) - K^* w^D (Ku - u_0) = 0, \tag{11}$$

or for the sake of the freezing-coefficient algorithm mentioned above, $L_z u = f$ with

$$f = K^* w^D u_0, \quad L_z = -\nabla \cdot z\nabla + K^* w^D K, \quad z = \frac{1}{|\nabla u|}. \tag{12}$$

The associated boundary condition along $\partial\Omega$ is again Neumann adiabatic.

What is the mathematical difference between the TV restoration models (5, 6) and the TV inpainting models, caused by the almost trivial modification of the weight function? Unlike the situation when the weight function $w(x) \geq \lambda/2 > 0$ for all pixels, the inpainting energy (10) is no longer *strictly* convex. Therefore, as shown by Chan, Kang, and Shen [**CKS01**], the existence of TV inpainting in $\mathrm{BV}(\Omega)$ is guaranteed, but generally uniqueness is not. In terms of the iterative algorithm based on the freezing-coefficient technique, the non-uniqueness is caused by the fact that the linearized operator L_z is only semi-positive definite.

As explained in our paper [**CKS01**], the non-uniqueness of TV inpainting may not be necessarily a defect of the model, but instead, an intrinsic part of the inpainting problem itself. To a certain degree, it models the uncertain situation of

human decision making when the given information is generated by two or more equally possible patterns.

All the computational methods discussed in Section 3.3 apply here too. In addition, to increase the sparsity of the linear system for general blurring kernel K, we have modified the freezing-coefficient iterative scheme at each step n to

$$[L_{z^{(n)}} + (w^D - K^* w^D K)]u^{(n+1)} = (w^D - K^* w^D K)u^{(n)} + f.$$

Figures 1 and 2 display the computational outputs for two images with simulated digital blurring, noising, and random packet loss with:

$$K_1 = \frac{1}{16^8}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}^{*8} \quad \text{and} \quad K_2 = \frac{1}{70^{30}}\begin{bmatrix} 1 & 3 & 1 \\ 20 & 20 & 20 \\ 1 & 3 & 1 \end{bmatrix}^{*30}. \tag{13}$$

Here the asterisks indicate that the powers are in the convolutional sense. K_1 and K_2 simulate the continuous isotropic blurring and horizontal motion blurring. These simulated results clearly demonstrate the power of TV inpainting for the potential market of noisy transmission of blurred images with randomly lost packets, images from the Hubble telescope, for example [**VO96, CW98**].

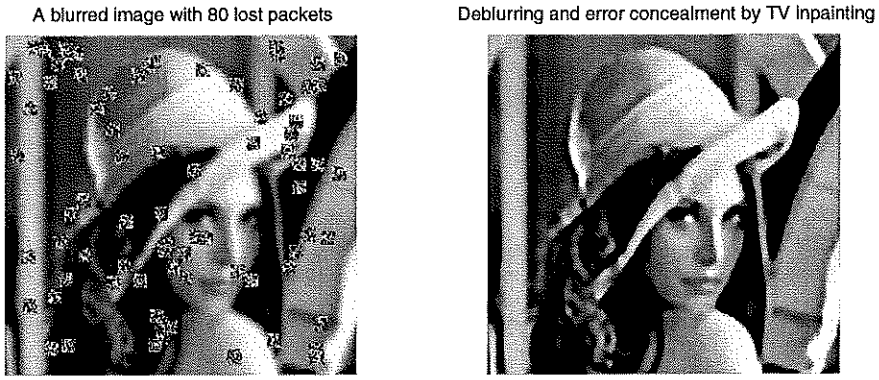

FIGURE 1. TV Inpainting of a blurred image (by K_1 in (13)) with simulated random loss of 80 packets.

There are also some important applications that cannot be covered by the continuous language, yet made possible by the extension of the TV norm onto general graphs by Chan, Osher, and Shen [**COS01**]. Zooming and perceptual image coding are such examples that we first studied in [**CS01a**].

Let us discuss a simplified version of the digital zoom-in problem and its TV inpainting approach. The goal of zoom-in is to create a $2N \times 2N$ digital image $[u_{i,j}]$ $(1 \le i, j \le 2N)$ from its possibly noisy coarse sampling $[u^0_{2i,2j}](1 \le i, j \le N)$.

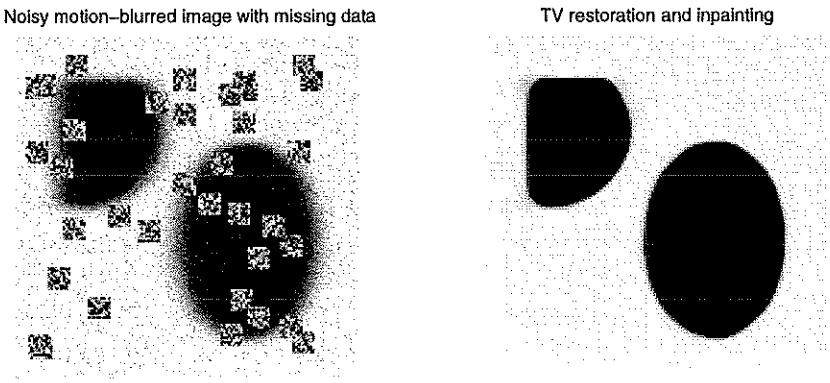

FIGURE 2. TV inpainting of a noisy blurred image (by the severe motion blurring K_2 in (13)) with simulated random packet loss.

Thus the weight function in this digital setting is given by

$$w^D_{i,j} = 1, \quad 2|i \ \& \ 2|j; \qquad 0, \quad \text{otherwise.}$$

In the paper [**CS01a**], we proposed the following zoom-in model based on TV inpainting:

$$\text{TV}_g([u_{i,j}]) + \sum_{1\le i,j\le 2N} (u_{i,j} - u^0_{i,j})^2 w^D_{i,j}, \tag{14}$$

where TV_g denotes the digital TV norm on graphs (a model for digital domains), as introduced in [**COS01**]. Figure 3 shows the computational output of the model when applied to a test image from Caltech's computational vision group. The comparison has been made between the BV image model and the Sobolev one (i.e., $E[u] = \int_\Omega |\nabla u|^2$). One can clearly observe that BV yields much better reconstruction in terms of the sharpness of boundaries.

The second non-trivial application connects inpainting to perceptual image coding, compression, and reconstruction [**CS01a**]. An example based on digital TV inpainting (14) (with a different weight w^D) is presented in Figure 4. We refer to our paper [**CS01a**] for more details.

These examples clearly demonstrate the beauty of a good image model - it facilitates all processing tasks, as a lighthouse does for successful navigation.

5. Beyond BV and Conclusion

5.1. The Mumford-Shah image model. A sibling to BV images is the celebrated object-edge model of Mumford and Shah [**MS89**]:

$$E_{\text{MS}}[u, \Gamma] = \int_{\Omega\setminus\Gamma} |\nabla u|^2 dx + \alpha H^1(\Gamma),$$

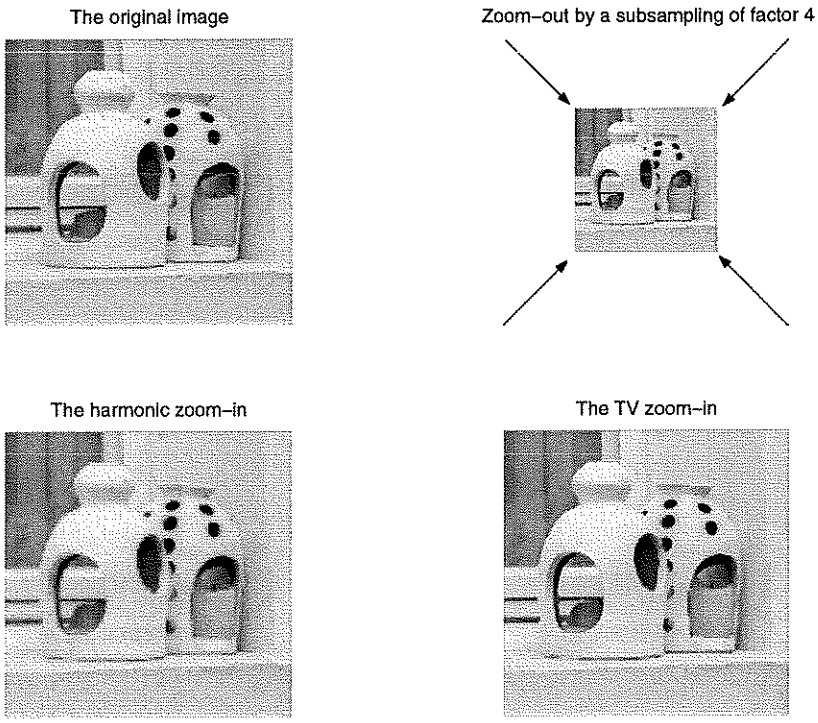

FIGURE 3. Bayesian zoom-in's based on the BV and Sobolev image models: TV inpainting yields much sharper boundaries [**CS01a**] (test image from Caltech's computational vision group).

where $H^1(\Gamma)$ is the one-dimensional Hausdorff measure of the "edge" set Γ, and $\alpha > 0$ is a fixed weight. It is easy to see that for near-cartoon images (i.e., the jump set Γ is piecewise smooth, and $|\nabla u| \ll 1$ on $\Omega \setminus \Gamma$), the Mumford-Shah image model $E_{\mathrm{MS}}[u, \Gamma]$ is equivalent to $\mathrm{TV}[u]$, since the latter as a Radon measure is concentrated along the jump set as well. In this case, the only major difference lies in that TV weights $H^1(\Gamma)$ adaptively along Γ based on the jump amplitude, while in E_{MS} the weight is uniformly fixed to be α.

The kinship between the two image models can also be seen from a unified viewpoint based on the edge signature function z introduced in Section 3.3. Such an approach has been well known in the vision community [**GR92**]. In Section 3.3, it has been established that the BV image model is approximately (controlled by $\epsilon \ll 1$) equivalent to

$$E^{\epsilon}_{\mathrm{TV}}[u, z] = \int_{\Omega} \left(z^2 |\nabla u|^2 + \epsilon z^2 + \frac{1}{z^2} \right) dx.$$

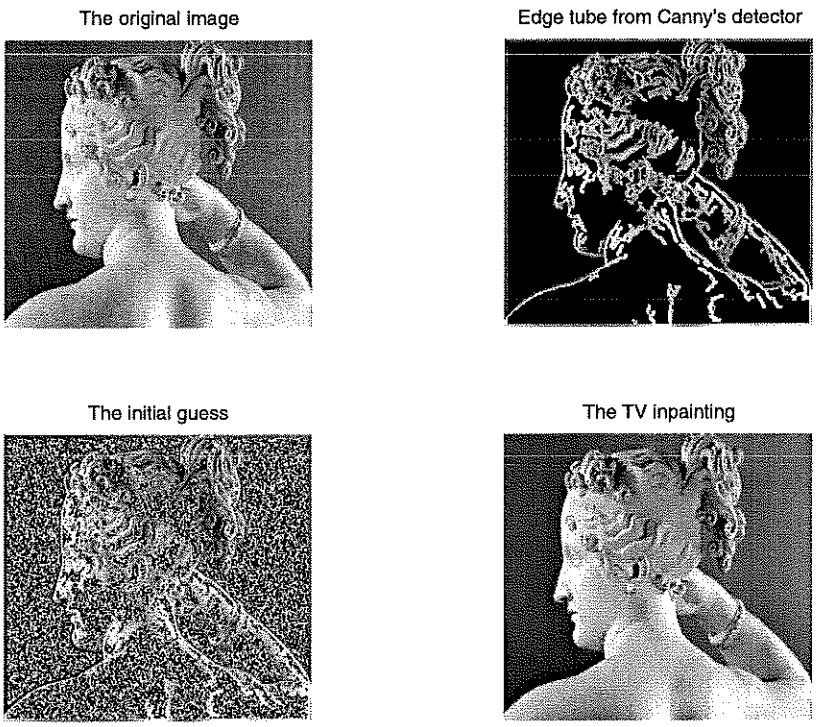

FIGURE 4. Perceptual image coding and decoding by the TV inpainting model [**CS01a**](test image from Caltech's computational vision group).

Here we have replaced the original z by z^2 since it is positive as seen from the freezing-coefficient algorithm in Section 3.3. On the other hand, under the Γ-convergence approximation theory (Ambrosio and Tortorelli [**AT90, AT92**]), the edge set Γ in the Mumford-Shah can also be replaced by an edge "signature" function z, and the image model is approximately equivalent to

$$E^{\epsilon}_{\mathrm{MS}}[u,z] = \int_{\Omega} \left(z^2|\nabla u|^2 + \alpha(\epsilon|\nabla z|^2 + \frac{(1-z)^2}{4\epsilon}) \right) \, dx.$$

Therefore, by introducing the edge signature z, both the BV image model of Rudin-Osher-Fatemi and the object-edge image model of Mumford and Shah belong to the same class of coupled energies:

$$E_{\epsilon}[u,z] = \int_{\Omega} (z^2|\nabla u|^2 + g_{\epsilon}(z, \nabla z, \nabla \otimes \nabla z)) \, dx,$$

where g_{ϵ} is a suitable function controlled by a small parameter ϵ, and $\nabla \otimes \nabla z$ denotes the Hessian.

The Mumford-Shah image model, once computationally realized (such as by the region growing method of Morel and Solimini [**MS95**], and by the level-set

method of Osher and Sethian [**OS88**], as recently studied by Tsai, Yezzi, and Willsky [**TAYW01**], and Chan and Vese [**CV01**]), is very powerful for image denoising and segmentation. Novel applications to the inpainting problem have been studied recently by Chan and Shen [**CS01a**], Tsai, Yezzi, and Willsky [**TAYW01**], and Esedoglu and Shen [**ES02**].

5.2. Gousseau and Morel: Natural images are NOT BV. This remarkable recent result of Gousseau and Morel [**GM01**] is the fruit of a successful combination of statistical image study and mathematical analysis of image models.

The key is the following lower bound for the TV norm by the so called *sectional density* $f_u(h,s)$:

$$\mathrm{TV}(u) \geq \pi^{\frac{1}{2}} \int_0^{|\Omega|} s^{\frac{1}{2}} f_u(h,s) ds,$$

where $|\Omega|$ denotes the Lebesgue measure (or area) of the image domain, s an "area" parameter, and h a quantization level of the image. Roughly speaking, for a given image u and a quantization level h, $f_u(h,s)ds$ denotes the number of disjoint h-"blobs" (or h-sections) whose areas fall within $[s, s+ds]$ (see [**GM01**] for more details). The empirical statistics based on many natural images, obtained by the same school of authors [**AGM99**], reveals the following power law:

$$f_u(h,s) = \frac{const.}{s^\alpha}, \quad \text{with} \quad \alpha \approx 2,$$

for any generic and homogeneous natural image u. Therefore,

$$\mathrm{TV}(u) \geq const. \int_0^{|\Omega|} \frac{1}{s^{\alpha-1/2}}\, ds,$$

which diverges at $s=0$ for all $\alpha \geq 3/2$.

In Meyer's $u+v$ language [**Mey01**], the result reveals that the v-component for generic natural images contains *too many* small scale "blobs" (clustering controlled by a quantization level h), which makes generally $u+v \notin \mathrm{BV}(\Omega)$.

Therefore, this negative assertion does not diminish the positive role of Rudin-Osher-Fatemi's BV images in the successful modeling of the u component.

5.3. Conclusion of the paper. A generic image seems to be the composition of two components: $u+v$, as Meyer [**Mey01**] puts it recently. Roughly speaking, u is the *deterministic* component, and v is the "texture" or "clutter" component [**MG01**], characterized by rapid oscillations but still away from being white noise. The v-component carries a delicate correlation between space and spatial frequencies, and as a result, statistical, spectral, and wavelets tools are ideal. The u-component is more geometric, and embedded with the crucial information of deterministic and large-scale features such as edges, corners, and T-junctions. The BV image model of Rudin, Osher and Fatemi is one of the very few successful

models for the u-component, which are both theoretically accessible and computationally efficient. For images with low textures (i.e., with negligible local ergodic variances), such as many indoor scenes capturing large objects, the BV image model by itself is often sufficient for the tasks like denoising, deblurring, and inpainting. This viewpoint has been strongly supported by various computational results.

Acknowledgments

We would like to thank all the people below for their generous teaching, inspiration, and support during this project (in a randomized order): Stan Osher, Luminita Vese, Sang Ha Kang, David Donoho, Li-Tien Cheng, Selim Esedoglu, Marcelo Bertalmio, Guillermo Sapiro, Dan Kerstan, Stu Geman, David Mumford, Jean-Michel Morel, Yann Gousseau, Rachid Deriche, Bob Gulliver, IMA and IPAM. We are also very grateful to all the authors mentioned in the paper who have been devoted to the understanding and development of the general mathematical foundation for image and vision analysis, especially the three recent publications by Meyer [**Mey01**], Gousseau and Morel [**GM01**], and Mumford and Gidas [**MG01**]. In addition, the second author thanks his dear friend and former advisor Gil Strang for his constant encouragement.

We dedicate our paper to the occasion of the 60th birthday of Stan Osher.

References

[AG01] A. Aldroubi and K. Gröchenig. Nonuniform sampling and reconstruction in shift-invariant spaces. *SIAM Review*, 43(4):585–620, 2001.

[AGM99] L. Alvarez, Y. Gousseau, and J.-M. Morel. The size of objects in natural and artificial images. *Advances in Imaging and Electron Physics*, 111:167–242, 1999.

[AT90] L. Ambrosio and V. M. Tortorelli. Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. *Comm. Pure Appl. Math.*, 43:999–1036, 1990.

[AT92] L. Ambrosio and V. M. Tortorelli. On the approximation of free discontinuity problems. *Boll. Un. Mat. Ital.*, 6-B:105–123, 1992.

[AV94] R. Acar and C. R. Vogel. Analysis of total variation penalty methods for ill-posed problems. *Inverse Prob.*, 10:1217–1229, 1994.

[BSCB00] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. Computer Graphics, *SIGGRAPH 2000*, July, 2000.

[BTU01] T. Blu, P. Thevenaz, and M. Unser. Moms: maximal-order interpolation of minimal support. *IEEE Trans. Image Process.*, 10(7):1069–1080, 2001.

[CD] E. J. Candés and D. L. Donoho. Curvelets and reconstruction of images from noisy radon data. *Wavelet Applications in Signal and Image Processing* VIII, A. Aldroubi, A. F. Laine, M. A. Unser eds., Proc. SPIE 4119, 2000.

[CGM99] T. F. Chan, G. H. Golub, and P. Mulet. A nonlinear primal-dual method for total variation-based image restoration. *SIAM J. Sci. Comput.*, 20(6):1964–1977, 1999.

[CKS01] T. F. Chan, S.-H. Kang, and J. Shen. Euler's elastica and curvature based inpaintings. UCLA CAM Report 2001-12 at: www.math.ucla.edu/~imagers; submitted to *SIAM J. Appl. Math.*, 2001.

[CL97] A. Chambolle and P. L. Lions. Image recovery via Total Variational minimization and related problems. *Numer. Math.*, 76:167–188, 1997.

[CM] T. F. Chan and P. Mulet. On the convergence of the lagged diffusivity fixed point method in total variation image restoration. *SIAM J. Numer. Anal.*, 36:354–367.

[COS01] T. F. Chan, S. Osher, and J. Shen. The digital TV filter and nonlinear denoising. *IEEE Trans. Image Process.*, 10(2):231–241, 2001.

[CS01a] T. F. Chan and J. Shen. Mathematical models for local non-texture inpaintings. *SIAM J. Appl. Math.*, 61(4):1019–1043, 2001.

[CS01b] T. F. Chan and J. Shen. Morphologically invariant PDE inpaintings. UCLA CAM Report 2001-15 at: www.math.ucla.edu/~imagers; submitted to *IEEE Trans. Image Process.*, 2001.

[CS01c] T. F. Chan and J. Shen. Nontexture inpainting by curvature driven diffusions (CDD). *J. Visual Comm. Image Rep.*, 12(4):436–449, 2001.

[CS01d] T. F. Chan and J. Shen. Variational restoration of non-flat image features: models and algorithms. *SIAM J. Appl. Math.*, 61(4):1338–1361, 2001.

[CS02] T. F. Chan and J. Shen. Bayesian inpainting based on geometric image models. Preprint, 2002.

[CV01] T. F. Chan and L. Vese. A level set algorithm for minimizing the Mumford-Shah functional in image processing. *IEEE/Computer Society Proceedings of the 1st IEEE Workshop on "Variational and Level Set Methods in Computer Vision"*, pages 161–168, 2001.

[CW98] T. F. Chan and C. K. Wong. Total variation blind deconvolution. *IEEE Trans. Image Process.*, 7(3):370–375, 1998.

[Don00] D. L. Donoho. Curvelets. Invited talk at workshop on *Wavelets, Statistics, and Image Processing*, Geogia Inst. Tech., 1999; Invited talk at MSRI workshop on *Mathematics of Imaging*; 1999. Beamlets, Invited talk at IMA workshop on *Image Analysis and Low Level Vision*, 2000.

[DS96] D. Dobson and F. Santosa. Recovery of blocky images from noisy and blurred data. *SIAM J. Appl. Math.*, 56(4):1181–1198, 1996.

[DV97] D. C. Dobson and C. R. Vogel. Convergence of an iterative method for total variation denoising. *SIAM J. Numer. Anal.*, 34(5):1779–1791, 1997.

[ES02] S. Esedoglu and J. Shen. Digital inpainting based on the Mumford-Shah-Euler image model. *European J. Appl. Math.*, in press, 2002.

[GG84] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.*, 6:721–741, 1984.

[Gib02] W. Gibbs. *Elementary Principles of Statistical Mechanics.* Yale University Press, 1902.

[Giu84] E. Giusti. *Minimal Surfaces and Functions of Bounded Variation.* Birkhäuser, Boston, 1984.

[GL83] G. H. Golub and C. F. Van Loan. *Matrix Computations.* The Johns Hopkins University Press, Baltimore, 1983.

[GM01] Y. Gousseau and J.-M. Morel. Are natural images of bounded variation. *SIAM J. Math. Anal.*, 33(3):634–648, 2001.

[GR92] D. Geman and G. Reynolds. Constrained image restoration and the recovery of discontinuities. *IEEE Trans. Pattern Anal. Machine Intell.*, 14:367–383, 1992.

[IP97] H. Igehy and L. Pereira. Image replacement through texture synthesis. *Proceedings of IEEE Int. Conf. Image Processing*, 1997.

[Kan79] G. Kanizsa. *Organization in Vision.* Praeger, New York, 1979.

[Ker] D. Kersten. High-level vision and statistical inference. In M. S. Gazzaniga, editor, *The New Cognitive Neurosciences*, pages 353–363. The MIT Press, Cambridge, MA, USA.

[LO01] X. Li and M.T. Orchard. New edge-directed interpolation. *IEEE Trans. Image Process.*, 10(10):1521 – 1527, 2001.

[Mey01] Y. Meyer. *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, volume 22 of *University Lecture Series.* AMS, Providence, 2001.

[MG01] D. Mumford and B. Gidas. Stochastic models for generic images. *Quaterly of Applied Mathematics*, 2001.

[MH80] D. Marr and E. Hildreth. Theory of edge detection. *Proc. Royal Soc. London*, B:207: 187–217, 1980.

[MO00] A. Marquina and S. Osher. Explicit algorithms for a new time dependent model based on level set motion for nonlinear deblurring and noise removal. *Siam. J. Sci. Comput.*, 22:387–405, 2000.

[MS89] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Applied. Math.*, 42:577–685, 1989.

[MS95] J.-M. Morel and S. Solimini. *Variational Methods in Image Segmentation*, volume 14 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser, Boston, 1995.

[Mum94] D. Mumford. *Geometry Driven Diffusion in Computer Vision*, chapter "The Bayesian rationale for energy functionals", pages 141–153. Kluwer Academic, 1994.

[NMS93] M. Nitzberg, D. Mumford, and T. Shiota. *Filtering, Segmentation, and Depth*. Lecture Notes in Comp. Sci., Vol. 662. Springer-Verlag, Berlin, 1993.

[OS88] S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.*, 79(12), 1988.

[PM00] E. L. Pennec and S. Mallat. Image compression with geometrical wavelets. *Proc. of IEEE ICIP*, 1:661–664, 2000.

[RO94] L. Rudin and S. Osher. Total variation based image restoration with free local constraints. *Proc. 1st IEEE ICIP*, 1:31–35, 1994.

[ROF92] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[Rud] L. Rudin. Images, numerical analysis of singularities, and shock filters. Caltech, Comp. Sci. Dept. Tech. Report 5250 , *1987*.

[SC] D. M. Strong and T. F. Chan. Edge-preserving and scale-dependent properties of Total Variation regularization. *SIAM Journal on Applied Mathematics*, submitted, 2000.

[Str93] G. Strang. *Introduction to Applied Mathematics*. Wellesley-Cambridge Press, MA, 1993.

[TAYW01] A. Tsai, Jr. A. Yezzi, and A. S. Willsky. Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation and magnification. *IEEE Trans. Image Process.*, 10(8):1169–1186, 2001.

[TD01] D. Tschumperle and R. Deriche. Regularization of orthonormal vector sets using coupled PDE's. *Proceedings of IEEE Workshop on Variational and Level Set Methods in Computer Vision*, pages 3–10, 2001.

[Tik63] A. N. Tikhonov. Regularization of incorrectly posed problems. *Soviet Math. Dokl.*, 4:1624–1627, 1963.

[TSC] B. Tang, G. Sapiro, and V. Caselles. Color image enhancement via chromaticity diffusion. Elect. Comp. Eng. Dept. Tech. Report, University of Minnesota, *1999*.

[VO96] C.R. Vogel and M. E. Oman. Iterative methods for total variation denoising. *SIAM J. Sci. Comput.*, 17(1):227–238, 1996.

[ZM97] S. C. Zhu and D. Mumford. Prior learning and Gibbs reaction-diffusion. *IEEE Trans. Pattern Anal. Machine Intell.*, 19(11):1236–1250, 1997.

[ZWM97] S. C. Zhu, Y. N. Wu, and D. Mumford. Minimax entropy principle and its applications to texture modeling. *Neural Computation*, 9:1627–1660, 1997.

Department of Mathematics, UCLA, Los Angeles, CA 90095, USA
E-mail address: TonyC@college.ucla.edu

School of Mathematics, (206 Church Street, SE,) University of Minnesota, Minneapolis, MN 55455, USA. Fax: (612) 626-2017. (Corresponding Author)
E-mail address: jhshen@math.umn.edu