SIAM J. APPL. MATH.
Vol. 62, No. 3, pp. 1019–1043

MATHEMATICAL MODELS FOR LOCAL NONTEXTURE INPAINTINGS*

TONY F. CHAN† AND JIANHONG SHEN‡

Dedicated to Stanley Osher on the occasion of his 60th birthday.

Abstract. Inspired by the recent work of Bertalmio et al. on digital inpaintings [SIGGRAPH 2000], we develop general mathematical models for local inpaintings of nontexture images. On smooth regions, inpaintings are connected to the harmonic and biharmonic extensions, and inpainting orders are analyzed. For inpaintings involving the recovery of edges, we study a variational model that is closely connected to the classical *total variation* (TV) denoising model of Rudin, Osher, and Fatemi [*Phys. D*, 60 (1992), pp. 259–268]. Other models are also discussed based on the Mumford–Shah regularity [*Comm. Pure Appl. Math.,* XLII (1989), pp. 577–685] and curvature driven diffusions (CDD) of Chan and Shen [*J. Visual Comm. Image Rep.*, 12 (2001)]. The broad applications of the inpainting models are demonstrated through restoring scratched old photos, disocclusion in vision analysis, text removal, digital zooming, and edge-based image coding.

Key words. inpainting, disocclusion, interpolation, variational/PDE method, prior image models, total variation, digital zooming, image coding

AMS subject classifications. Primary, 94A08; Secondary, 68U10, 65K10

PII. S0036139900368844

1. Introduction.

To inpaint a damaged image or an ancient painting with missing regions is to guess and fill in the lost image information in such a consistent way that the restored image or painting seems as natural as its original version. (See Figure 1.1.)

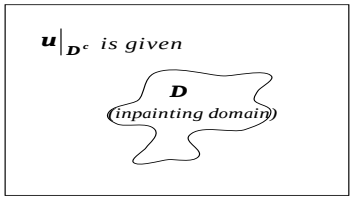

FIG. 1.1. *Inpainting is to paint the missing* $u|_D$ *on an inpainting domain* D *based on the image information available outside.*

The word "inpainting" was initially invented by museum or art restoration workers [17, 47]. The concept of *digital inpainting* was only recently introduced into digital image processing by Bertalmio et al. [1], who were the first group to develop inpainting models based on high order PDEs. Earlier related works based on second order PDEs and variational techniques can be found in Caselles, Morel, and Sbert [4] and Masnou and Morel [31]. Meanwhile, in computer science, much research similar to

*Received by the editors March 10, 2000; accepted for publication (in revised form) May 9, 2001; published electronically February 6, 2002. This research was supported by grants from the NSF under grant number DMS-9973341 and from the ONR under N00014-96-1-0277.
http://www.siam.org/journals/siap/62-3/36884.html
†Institute of Pure and Applied Mathematics (IPAM), UCLA, Los Angeles, CA 90095-7121 (chan@ipam.ucla.edu).
‡School of Mathematics, University of Minnesota, Minneapolis, MN 55455 (jhshen@math.umn.edu).

the inpainting problem has also been carried out in the context of image interpolation [26], image replacement [22], and error concealment [23, 27], although these works are more based on statistical and algorithmic approaches.

Important applications of digital inpainting include (a) digital restoration of ancient paintings for conservation purposes [17, 47], (b) restoring aged or damaged photographs and films [25, 26], (c) text removal and object removal in images for special effects [1], (d) disocclusion in vision research [31, 36], (e) digital zooming and edge-based image coding (sections 8 and 9).

Mathematically, what makes the inpainting problem so challenging is the complexity of image functions. Unlike many traditional interpolation or boundary value problems, the target image functions to be inpainted typically lie outside the Sobolev category. Some examples include: (a) natural images (*clutters*) are modeled by distributions (Mumford [33]); (b) texture images contain very rich statistical content and are modeled by Markov random fields and Gibbs fields (Geman and Geman [18], Brémaud [3]); and (c) most nontexture images can be well approximated by functions with bounded variations (Rudin and Osher [40], Rudin, Osher, and Fatemi [41], Chambolle and Lions [5]) and the celebrated Mumford and Shah object-boundary model [34]. Such multilevel complexities of image functions force researchers to develop inpainting schemes targeted at specific classes of images. As a result, these inpainting models are of low levels. The ultimate goal, of course, as in the blueprint of vision and artificial intelligence, is eventually to be able to combine and integrate all the low-level inpainting components into an ideal program that can well approximate human inpainters.

The current paper represents a first systematic step toward this goal, and the restrictive words "local" and "nontexture" in the title clearly indicate the low-level nature of all inpainting models developed in this paper. The crucial concept and principle of "locality" will be explained in the next section. And the reason that the current paper does not touch texture inpainting is that all inpainting models here are based on the variational principle or PDEs, and the resulting regularity requirement is unsuitable for general statistical textures. For inpaintings of textures, some recent work has been accomplished by Wei and Levoy [48] and Igehy and Pereira [22].

The paper is organized as follows. Section 2 clarifies the meaning of locality through two examples connected to human visual inference. In section 3, we study inpainting models and their accuracy analysis for smooth images. The key tool is Green's second formula, which leads to linear and cubic schemes realized by *harmonic* and *biharmonic* inpaintings. In section 4, we propose three *inpainting principles* for a realistic low-level inpainting model. In this spirit, the *total variation* (TV) inpainting model is formulated in section 5, which extends the classical TV denoising model of Rudin and Osher [40] and Rudin, Osher, and Fatemi [41]. The digital implementation of the TV inpainting model is also presented. In section 6, we propose a segmentation-based inpainting scheme, as inspired by the well-known image model of Mumford and Shah [34]. Section 7 introduces the so-called *connectivity principle* and the recent model of Chan and Shen [11] on inpaintings based on *curvature driven diffusions* (CDD). In section 8, for the first time in the literature of image processing, we make the link between *digital zoom-ins* and TV inpaintings. A digital zoom-in model almost identical to the continuous TV inpainting model is constructed based on the self-contained digitized PDE method developed by Chan, Osher, and Shen [9]. Section 9 explains another new important application of the inpainting technique to edge-based image coding schemes. The last section demonstrates many interesting applications of the inpainting models developed in the paper.

2. Why local inpaintings. We discuss the concept of locality for low-level inpainting models from the viewpoint of human visual inference.

Local inference. First, the locality condition means that the models do not rely on global feature or pattern recognition. Inpainting is thus based only on local (with respect to the inpainting domain) information. For nontexture images, a pattern is a class of spatial or color symmetry, such as mirror symmetry, translation invariance, and periodicity. For example, most human faces are almost mirror symmetric along the mouth-nose central line. Though apparent to human visual perception, such patterns are much more difficult and costly to catch via digital intelligence, due to the rich variation in scales and structures for general images.

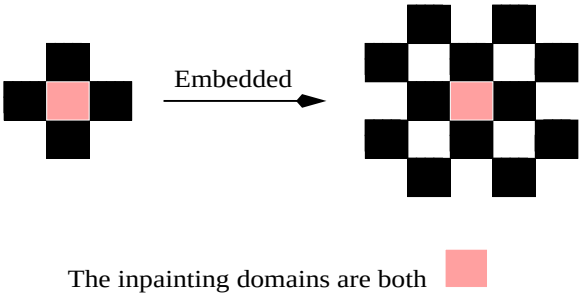

FIG. 2.1. *A local inpainting scheme does not require global pattern recognition.*

A classical example in vision analysis as shown in Figure 2.1 can clarify the above discussion. For the image to be inpainted on the left, the inpainting (or occluded) domain is the gray square at the intersection. Human observers can usually easily "see" a complete black cross and thus fill in the black color. Most of us would agree that it gives the best guess. In the right panel, the image on the left is embedded into a larger structure. One can easily recognize the global chessboard pattern and thus fill in the white color to complete the spatial symmetry. Therefore, human perceptual inference depends on the global context.

Such complexity of human visual inference parallels that of inpainting models. Any high-level inpainting scheme must be able to carry out pattern recognition. In this paper, the inpainting models should be considered as low-level ones—the inpainting outputs are independent of global patterns. Therefore, even for the right panel in Figure 2.1, the inpainted color for the missing square domain will still be black.

The factor of scale or aspect ratio. Scale plays a universally significant role in image and vision analysis. Thus it also does in the problem of inpainting.

Consider Figure 2.2. In the left panel, the inpainting scale L is much larger than that of the characteristic feature (denoted by l), and the left part "E" and right part "3" seem to be more uncorrelated. We thus tend to accept the figure as two separated letters "E 3." The image on the right, on the other hand, has an inpainting scale L smaller than l. Accordingly, we are more likely to believe that the figure is a broken letter "B." In this example, the nonuniqueness is not caused by global patterns, but by our guess on the correlation among features left there. The controlling parameter is thus the *scale* or *aspect ratio*. The TV inpainting model and the segmentation-based inpainting model developed later can both imitate this effect. However, as discussed in section 7, for many applications in image processing, due to the large dynamic range of scales present, the *connectivity principle* must be enforced regardless of the scale factor. Therefore, the major inpainting models developed in the current paper

FIG. 2.2. *The effect of the inpainting scale L.*

see their best performance in inpainting problems whose inpainting domains are small or local, as in the right panel of Figure 2.2.

3. Inpaintings of smooth images and Green's second formula. To develop a rigorous mathematical framework for inpaintings, as is well practiced in numerical analysis, we start from a simple setting, in which the accuracy of inpainting can be well studied. This is the case when the target image functions are smooth, or the inpainting domains are contained in the interior of smooth two-dimensional (2-D) objects. This simple model serves as the first step toward more general and realistic inpainting models.

Let u^0 be a smooth image function defined on a 2-D domain Ω (a rectangular domain, typically). Denote the domain to be inpainted by D, its diameter by d, and the restriction of u^0 on D by $u^0\big|_D$. Then to inpaint is to construct a good approximation u_D to $u^0\big|_D$.

An inpainting scheme is said to be *linear* if for any smooth test image u^0, as the diameter d of the inpainting region D shrinks to 0,

$$\|u_D - u^0\big|_D\|_\infty = O(d^2). \tag{3.1}$$

Similarly, an inpainting scheme is said to be of kth order if

$$\|u_D - u^0\big|_D\|_\infty = O(d^{k+1}). \tag{3.2}$$

3.1. Smooth inpainting via Green's second formula. Recall that for the 1-dimensional (1-D) case, harmonic functions on an interval have to be linear functions. Therefore, 1-D linear inpaintings may be carried out equivalently by harmonic extensions, which provide the key to 2-D smooth inpaintings. Here we propose to apply Green's second formula.

Let Δ denote the Laplacian

$$\Delta u := \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}.$$

Then Green's second formula on D is

$$\int_D (u\Delta v - v\Delta u)dxdy = \int_\Gamma \left(u\frac{\partial v}{\partial \mathbf{n}} - v\frac{\partial u}{\partial \mathbf{n}}\right) ds, \tag{3.3}$$

where

(a) u and v are any C^2 functions defined on the closure of D;

(b) $\mathbf{n}$ is the outward (w.r.t. D) normal direction of Γ, and s the length parameter.

Take $G(z_0, z)$ to be the Green's function for the grounded Poisson equation on D. That is, for any "source" point $z_0 = (x_0, y_0) \in D$, as a function of the "field" point $z = (x, y) \in D$, $G(z_0, z)$ solves

$$-\Delta G = \delta(z - z_0), \qquad G\big|_\Gamma = 0.$$

Applying Green's second formula to $(u = u^0(z), v = -G(z_0, z))$, we have

$$u^0(z_0) = \int_\Gamma u^0(z(s)) \frac{\partial(-G(z_0, z))}{\partial \mathbf{n}} ds + \int_D G(z_0, z) \left(-\Delta u^0(z)\right) dz, \tag{3.4}$$

where $dz = dxdy$. (More rigorously, we should have used the symbol $(dz \wedge d\overline{z})/2i$.)

In (3.4), denote the first term on the right by $u^h(z_0)$, and the second term by $u^a(z_0)$. Then u^h is the harmonic extension of $f = u^0\big|_\Gamma$, and

$$d\omega_{z_0} = \frac{\partial(-G(z_0, z))}{\partial \mathbf{n}} ds$$

is the harmonic measure of Γ associated with a source point z_0 (Nevanlinna [35]).

The antiharmonic component $u^a := u^0 - u^h$ satisfies the Poisson equation

$$\Delta u^a(z) = \Delta u^0(z), \quad z \in D, \qquad \text{and} \quad u^a\big|_\Gamma = 0. \tag{3.5}$$

Computationally, the Poisson equation is favored over the direct integration formulation since one can profit from many numerical PDE schemes and their fast solvers.

To establish a rigorous result on the inpainting accuracy for smooth images, we turn to the geometry of a 2-D domain encoded into its associated Green's function. The following results on Green's functions are indeed standard. We include them here due to the increasingly important role played by the complex potential theory in signal and image processing. (For example, recent applications of the complex potential theory to digital signal processing have been studied in [42, 43].)

THEOREM 3.1. *Let d denote the diameter of a domain D and $G(z_0, z)$ the associated Green's function for the Poisson equation. Then*

$$\int_D G(z_0, z) dxdy \le \frac{d^2}{4}.$$

The proof is based upon two simple lemmas.

LEMMA 3.2 (comparison lemma). *Suppose $D_1 \subset D_2$, and $G_1(z_0, z)$ and $G_2(z_0, z)$ are their associated Green's functions. Then for all $z_0, z \in D_1$,*

$$G_1(z_0, z) \le G_2(z_0, z).$$

Proof. For any $z_0 \in D_1$, define

$$g(z) = G_2(z_0, z) - G_1(z_0, z).$$

Then along the boundary of D_1,

$$g(z) = G_2(z_0, z) \ge 0,$$

since the grounded Green's function is always nonnegative. Moreover, $g(z)$ is harmonic inside D_1 because the logarithm singularities at z_0 are canceled out. Therefore $g(z) \ge$

0 for all $z \in D_1$ due to the *maximum principle* of harmonic functions: The minimum is always achieved along the boundary (Gilbarg and Trudinger [19]). This proves the lemma. □

LEMMA 3.3. *Suppose B_1 is the unit disk centered at 0, and $G_1(z_0, z)$ its Green's function. Then*

$$\int_{B_1} G_1(z_0, z)dxdy = \frac{1 - |z_0|^2}{4}$$

for all $z_0 \in B_1$.

Proof. Consider the Poisson equation on B_1

$$-\Delta u = 1, \qquad u\big|_{\partial B_1} = 0.$$

It is easy to see that the unique solution is

$$u(z) = \frac{1 - |z|^2}{4} = \frac{1 - x^2 - y^2}{4}.$$

On the other hand, by Green's second formula,

$$u(z_0) = \int_{B_1} G_1(z_0, z)(-\Delta u(z))dxdy = \int_{B_1} G_1(z_0, z)dxdy.$$

This verifies the lemma. (Note: Since we do know that

$$G_1(z_0, z) = \frac{-1}{2\pi} \ln \left| \frac{z - z_0}{1 - \overline{z_0} z} \right|,$$

the lemma can also be worked out by evaluating the integral explicitly.) □

We are now ready to give a proof of Theorem 3.1.

Proof. Take any single point $w \in D$, and let B_d denote the disk centered at w and with radius d. Then

$$D \subset B_d.$$

Let $G_d(z_0, z)$ denote the Green's function for B_d. Then Lemma 3.2 shows that

$$G(z_0, z) \le G_d(z_0, z),$$

for all z_0 and z in D. For simplicity, let us assume that $w = 0$. Then we have the scaling law

$$G_d(z_0, z) = G_1\left(\frac{z_0}{d}, \frac{z}{d}\right), \tag{3.6}$$

where G_1, as in Lemma 3.3, is the Green's function for B_1. (This scaling law is true only for the 2-D case.) Therefore, by Lemma 3.3, for any $z_0 \in D$,

$$\begin{aligned}
\int_D G(z_0, z)dxdy &\le \int_D G_d(z_0, z)dxdy \le \int_{B_d} G_d(z_0, z)dxdy \\
&= \int_{B_d} G_1\left(\frac{z_0}{d}, \frac{z}{d}\right) dxdy = d^2 \int_{B_1} G_1\left(\frac{z_0}{d}, z'\right) dx'dy' \\
&= d^2 \frac{1 - |z_0/d|^2}{4} \le \frac{d^2}{4},
\end{aligned}$$

as asserted by the theorem. (The last step is due to our assumption that $w = 0 \in D$ and $z_0 \in D$. If this is not the case, then simply replace z_0 and z by $z_0 - w$ and $z - w$, and the proof still holds.) This completes the proof. □

Based on this theorem, we can easily establish the accuracy orders for inpaintings based on Green's second formula.

(a) *Linear inpainting via harmonic extension.* Suppose we inpaint $u^0\big|_D$ simply by the harmonic extension, i.e., $u_D = u^h$. We now show that this is a linear inpainting scheme, i.e.,

$$\|u^h - u^0\big|_D\|_\infty = O(d^2)$$

as the diameter $d \to 0$.

According to (3.4), the error of the harmonic inpainting is exactly the anti-harmonic component u^a. Since u^0 is a fixed smooth function, there exists a constant M such that

$$|\Delta u^0(z)| \le M$$

for all $z \in D$. Then for any $z_0 \in D$, by Theorem 3.1,

$$|u^a(z_0)| \le M \int_D G(z_0, z)dz \le \frac{Md^2}{4}.$$

This validates the assertion.

(b) *Cubic inpainting via Green's formula.* To improve the accuracy, we must also inpaint the "detail" component u^a missed by the harmonic inpainting. (This idea is very close to the multiresolution synthesis in wavelet decomposition or coding [14, 44].)

Let u_D^Δ be *any* linear inpainting of $\Delta u^0\big|_D$ (via the harmonic scheme, for example). Then we inpaint $u^a\big|_D$ by u_D^a according to the integration formula

$$u_D^a(z_0) = \int_D G(z_0, z)(-u_D^\Delta(z))dz, \tag{3.7}$$

or equivalently, by solving the grounded Poisson equation

$$-\Delta u_D^a(z) = -u_D^\Delta(z), \quad z \in D; \qquad u_D^a\big|_\Gamma = 0.$$

Finally, by adding this new detail to the harmonic inpainting, we derive a more accurate inpainting u_D to the original smooth test image u^0:

$$u_D(z) = u^h(z) + u_D^a(z). \tag{3.8}$$

THEOREM 3.4 (cubic inpainting). *If u_D^Δ is a linear inpainting of Δu^0 on D, then* (3.7) *and* (3.8) *define a* cubic *inpainting of u^0, i.e.,*

$$\|u_D - u^0\big|_D\|_\infty = O(d^4).$$

Proof. By Green's second formula, for any $z_0 \in D$,

$$u_D(z_0) - u^0\big|_D(z_0) = \int_D G(z_0, z)\left(-u_D^\Delta(z) + \Delta u^0(z)\right)dxdy.$$

Since $u_D^{\Delta}(z)$ is a linear inpainting of $\Delta u^0(z)$, there exists a constant M, independent of the inpainting domain D, such that

$$|u_D^{\Delta}(z) - \Delta u^0(z)| \leq Md^2$$

for all $z \in D$. Hence,

$$|u_D(z_0) - u^0\big|_D(z_0)| \leq Md^2 \int_D G(z_0, z)dxdy.$$

The proof is then complete by Theorem 3.1. □

Remark 1. In the above cubic inpainting process, if the linear inpainting u_D^{Δ} of $\Delta u^0\big|_D$ is realized by the harmonic inpainting, then the cubic inpainting is in fact a biharmonic inpainting. That is, $u_D(z)$ solves the following biharmonic boundary value problem:

$$\Delta^2 u_D = 0, \qquad u_D\big|_\Gamma = u^0\big|_\Gamma, \qquad \Delta u_D\big|_\Gamma = \Delta u^0\big|_\Gamma.$$

4. Three principles for a practical inpainting scheme. As in the classical approximation theory, the smooth inpainting models have allowed us to study rigorously the inpainting accuracies. They also shed some light on the nature of the inpainting problem. In most applications, however, such models are less practical for the following reasons:

(a) Images (even nontexture ones) are *deterministically* not smooth functions. They contain edges and discontinuities (see Figure 4.1).

(b) Images are often *statistically* corrupted by noise.

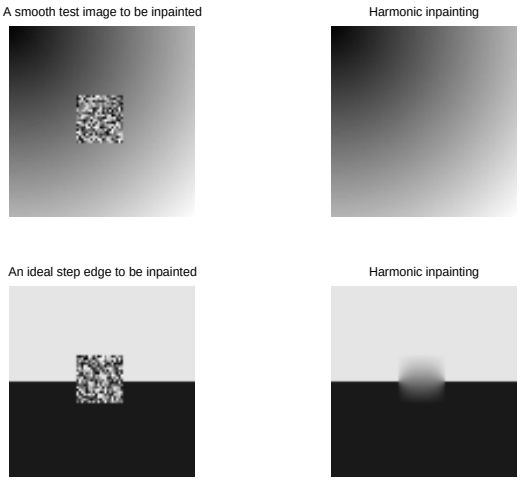

FIG. 4.1. *Harmonic inpaintings of a smooth image* ($u = r = \sqrt{x^2 + y^2}$) *and an ideal step edge.*

Human inpainters seem to have no difficulty in dealing with these two factors, and we intend to design more realistic low-level inpainting models that can at least imitate such functions. Thus we propose the following three inpainting principles for the next step of model construction.

(a) *Inpainting principle* I. The model shall be *local*. Since we restrict ourselves to models which do not require global learning, the inpainting u_D must be

completely determined by the existing information of u^0 in the vicinity of the inpainting domain D.

(b) *Inpainting principle* II. The model must be able to *restore narrow broken smooth edges*. We must take care of edge inpainting, since edges are crucial for object recognition and image segmentation, and in most practical examples they are indeed broken or occluded due to the large dynamic range of scales in images. However, generally, we shall not expect to restore widely broken edges because of the scale factor discussed in section 2.

(c) *Inpainting principle* III. The model must be *robust to noise*. This is because to human vision it is an easy task (when the noise is below a reasonable level) to detect clean features from the existing noisy image data and then extend them into the inpainting domain.

Both the linear harmonic inpainting and cubic biharmonic inpainting are models for *smooth* images. Hence they do not fit with inpainting principles II and III, although they are indeed *local* since only the behavior of u^0 near a neighborhood of Γ is needed (for digitally obtaining the traces of u^0 and Δu^0 along Γ).

In what follows, we shall study inpainting models that conform to these three principles.

5. The total variation (TV) inpainting model.

5.1. Formulation of the TV inpainting model. Let D be an inpainting domain with piecewise smooth boundary Γ, and E any fixed closed domain in the complement D^c, so that Γ lies in the *interior* of $E \cup D$ (Figure 5.1). Such a setting is motivated by the inpainting principles on locality and robustness to noise.

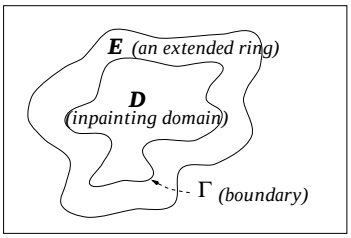

FIG. 5.1. *The TV inpainting model finds the best guess for* $u\big|_D$ *based on the TV norm on the extended domain* $E \cup D$ *and the noise constraint on* E.

We assume that $u^0\big|_E$ is contaminated by homogeneous white noise (modeled by the Gaussian distribution). The variational inpainting model is to find a function u on the extended inpainting domain $E \cup D$ such that it minimizes an appropriate regularity functional,

$$R[u] := \int_{E\cup D} r(|\nabla u|)dxdy, \tag{5.1}$$

under the fitting (or denoising) constraint on E

$$\frac{1}{\text{Area}(E)} \int_E |u - u^0|^2 dxdy = \sigma^2. \tag{5.2}$$

Here,

(i) r is an appropriate real function which is nonnegative for nonnegative inputs;

(ii) σ is the standard deviation of the white noise.

Remark 2. If D is empty, the above variational formulation belongs to the classical denoising models such as the H^1 model if $r(s) = s^2$, and the total variation model of Rudin, Osher, and Fatemi [41] if $r(s) = s$.

The variational formulation (5.1) plus (5.2) has been designed to satisfy inpainting principle I on locality and inpainting principle III on robustness to noise. To meet the second principle on the capability of restoring broken edges, we need to choose an appropriate regularity functional $R[u]$ or $r(s)$. Along a step edge, ∇u is a 1-D delta function δ_1 (like $\delta(x)$ as a function of x and y). Thus, to be able to restore a broken step edge, we have to require

$$\int_{E\cup D} r(\delta_1)dxdy$$

to be finite. This implies that if

$$r(s) = s^\alpha + (\text{lower order terms})$$

for some power α as $s \to +\infty$, then $\alpha \leq 1$. To ensure near convexity, $\alpha = 1$ is the ideal choice. It leads to the well-known TV restoration model of Rudin, Osher, and Fatemi [41], where $r(s)$ is taken to be s exactly. In this paper, we shall make the same choice, and call the inpainting model the *TV inpainting model*. It conforms to all three inpainting principles.

As practiced in the variational methodology [34, 41], it is more convenient to solve the unconstrained TV inpainting problem

$$J_\lambda[u] = \int_{E\cup D} |\nabla u| dxdy + \frac{\lambda}{2}\int_E |u - u^0|^2 dxdy, \tag{5.3}$$

where λ plays the role of the Lagrange multiplier for the constrained variational problem (5.1)–(5.2).

The Euler–Lagrange equation for the energy functional J_λ is

$$-\nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right) + \lambda_e(u - u^0) = 0 \tag{5.4}$$

for all $z = (x, y) \in E \cup D$, plus the Neumann boundary condition [10, 41]. Here the extended Lagrange multiplier λ_e is given by

$$\lambda_e = \begin{cases} \lambda, & z \in E, \\ 0, & z \in D. \end{cases}$$

The infinitesimal steepest descent equation for $J_\lambda[u]$ is therefore given by

$$\frac{\partial u}{\partial t} = \nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right) + \lambda_e(u^0 - u). \tag{5.5}$$

Since λ_e takes two different values, (5.4) or (5.5) is a two-phase problem, and the interface is the boundary Γ of the inpainting domain.

From the numerical point of view, in all of the above differential equations we replace the curvature term

$$\nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right) \qquad \text{by} \qquad \nabla\cdot\left(\frac{\nabla u}{|\nabla u|_a}\right), \tag{5.6}$$

where the "lifted" absolute value is defined by

$$|s|_a := \sqrt{s^2 + a^2}$$

for some (usually small) positive lifting parameter a. This corresponds to the choice of $r(s) = \sqrt{s^2 + a^2}$ for the regularizer $R[u]$ in (5.1). We are thus actually minimizing

$$J^a_\lambda[u] = \int_{E\cup D} \sqrt{a^2 + |\nabla u|^2}\, dxdy + \frac{\lambda}{2}\int_E |u - u^0|^2 dxdy.$$

As in most processing tasks involving thresholdings (like denoising and edge detection), the lifting parameter a also plays a thresholding role. In smooth regions where $|\nabla u| \ll a$, the model tries to imitate the harmonic inpainting, while along edges where $|\nabla u| \gg a$, the model resumes the TV inpainting.

On the other hand, from the theoretical point of view, the lifting parameter a also better conditions the TV inpainting model (5.3). In a noise-free situation, (5.3) is reduced to a boundary value problem:

$$\nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right) = 0, \quad x \in D; \qquad u\big|_{\partial D} = u^0\big|_{\partial D}. \tag{5.7}$$

As explained in [4], this boundary value problem, unlike harmonic extensions, is generally ill-posed and may fail to have or to *uniquely* have a solution. The parameter a plays a conditioning role as follows. For the lifted model,

$$\nabla \cdot \left(\frac{\nabla u}{|\nabla u|_a}\right) = \frac{1}{|\nabla u|^3_a}\left(|u_y|^2_a u_{xx} + |u_x|^2_a u_{yy} - 2u_x u_y u_{xy}\right).$$

As a second order equation, its local symbol σ_a is

$$\sigma_a = \frac{1}{|\nabla u|^3_a}\begin{bmatrix} |u_y|^2_a & -u_x u_y \\ -u_x u_y & |u_x|^2_a \end{bmatrix} = \frac{a^2}{|\nabla u|^3_a} I_2 + \frac{|\nabla u|^3}{|\nabla u|^3_a}\sigma_0,$$

where σ_0 is the symbol for the original TV model. Then it is easy to show the following.

PROPOSITION 5.1 (the conditioning effect of a). *The TV symbol σ_0 has eigenvalues 0 and $|\nabla u|^{-1}$, while the lifted TV symbol σ_a satisfies*

$$\frac{|\nabla u/a|_1^{-3}}{a} I_2 \le \sigma_a \le \frac{2}{a} I_2.$$

Therefore, at each pixel away from the edges (where $|\nabla u|$ is finite), the lifted TV equation is strongly elliptic; if u has a bounded gradient, then the lifted TV equation is in fact uniformly strongly elliptic. This is the conditioning effect of a.

Remark 3. In the most recent work of Chan, Kang, and Shen [6], the existence of solutions to the variational TV inpainting model (5.3) has been established in the space of functions with bounded variations. The issue of uniqueness is also discussed from the vision research point of view.

Remark 4. So far, the TV inpainting model has been solely motivated by the three inpainting principles. We now further justify the TV inpainting model through a well-known class of illusions in visual perception.

The vision phenomenon we are to discuss is best illustrated through the example of *Kanizsa's entangled woman and man*, which is one of the many artistic inventions

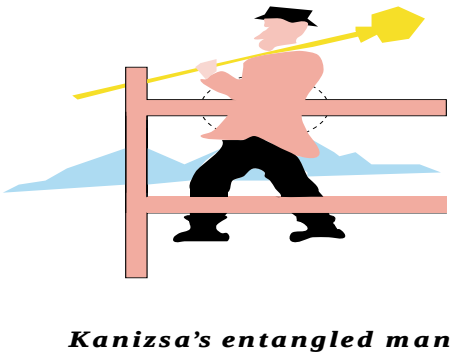

FIG. 5.2. *Can the TV inpainting model explain Kanizsa's entangled man?*

of Kanizsa [24]. Its importance for the mathematical understanding and modeling of human vision was first emphasized in Nitzberg, Mumford, and Shiota's systematic work on disocclusion [36]. We have plotted a simplified version in Figure 5.2, which we call "Kanizsa's entangled man."

Figure 5.2 shows how our visual perception can subconsciously contradict common knowledge in life. What we perceive is a man entangled in the fence. Knowing by common sense that he is behind the fence does not erase this false perception. As Nitzberg, Mumford, and Shiota [36] wrote, "Simply put, we navigate in the world successfully by seeing what's in front of what independently of knowing what's what." We now apply the TV inpainting model to explain such a "stubborn best guess" by our visual perception.

The contradiction occurs inside the circled region in Figure 5.2: the "fact" is that the upper body of the man is behind the fence, while our perception strongly prefers the opposite scenario. This disjunction is apparently caused by the presence of a color shared by the fence and the man's upper body. So the puzzle is, Why does human perception prefer to assign the controversial intersection to the upper body?

Kanisza's original explanation was based on the modal and amodal completion accomplished by the shortest edge continuation between T-junctions. Here we show that the TV inpainting model offers another similar explanation. While in practice the detection of T-junctions often relies on the sharpness of edges, our functional approach based on the variational principle seems to be more general.

First we simplify the problem to that of the left image in Figure 5.3. The vertical and horizontal bars separately model the man's upper body and the fence. Notice the length scales $L > l$; in Figure 5.2, L is roughly a triple of l. Assume that the two bars share the same gray level $u_b = u_f = 1/2$ (with "b" and "f" tracking the "body" and "fence" variables). The uncertain region is denoted by D.

Outside D, let us make a small perturbation of the two gray levels:

$$u_b = 1/2 \to u_b = 1/2 + \epsilon, \qquad u_f \to u_f = 1/2 - \epsilon$$

for some small positive gray value ϵ (see the image in the right panel of Figure 5.3). Now treat D as an inpainting domain and denote by u_D the optimal solution on D obtained from the TV inpainting model with $\lambda = \infty$ (since there is no noise) and E the complement of D. A simple calculation shows that

$$u_D = u_b = 1/2 + \epsilon, \tag{5.8}$$

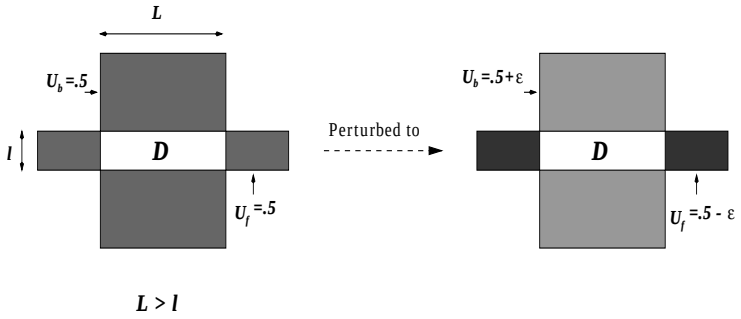

FIG. 5.3. *The model for Kanizsa's entangled man.*

which coincides with our "stubborn" perception. In other words, the TV model is consistent with the "algorithm" performed by our visual neurons.

In fact, it is easy to see that the optimal solution u_D must be a constant, say c. Then the maximum principle [9] requires that $u_f \leq c \leq u_b$. The total variation of u_D on the closure of D concentrates along the four edges and equals (Giusti [20])

$$2 \times (|u_f - c| * l + |u_b - c| * L) = [(1 + 2\epsilon)L - (1 - 2\epsilon)l] - (L - l)c. \tag{5.9}$$

We do not care about the TV measure on E because it is a fixed quantity for this noise-free inpainting problem. To minimize the TV norm as given in (5.9), the only choice is $c = u_b = 1/2 + \epsilon$, since $L > l$. This proves the claim.

5.2. Numerical implementation. If the inpainting domain D is empty, then (5.4) and (5.5) together comprise exactly the Rudin–Osher–Fatemi [41] denoising and deblurring restoration model. Its theoretical study can be found in Chambolle and Lions [5]. Numerical investigations and discussions can be found in [2, 8, 15, 41], and more recent ones in [9, 29, 38]. New applications of the TV model for restoring nonflat image features such as optical flows and chromaticity have appeared in the recent works of Perona [39]; Tang, Sapiro, and Caselles [45, 46]; Chan, Kang, and Shen [7]; and Chan and Shen [10].

In this paper, we have adopted the following numerical scheme for the TV inpainting model (5.4). Here we look for the steady solution directly, instead of by time marching (5.5), which is usually slow due to the time step constraints imposed by numerical stability.

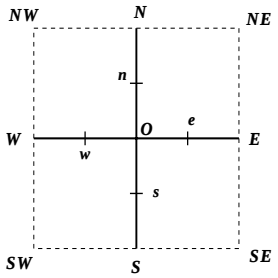

FIG. 5.4. *A target pixel O and its neighbors.*

As in Figure 5.4, at a given target pixel O, let E, N, W, S denote its four adjacent pixels, and e, n, w, s the corresponding four midway points (not directly available from

the digital image). Write

$$\Lambda_O = \{E, N, W, S\}.$$

Let $\mathbf{v} = (v^1, v^2) = \nabla u/|\nabla u|$. Then the divergence is first discretized by central differencing:

$$\nabla \cdot \mathbf{v} = \frac{\partial v^1}{\partial x} + \frac{\partial v^2}{\partial y} \tag{5.10}$$

$$\simeq \frac{v_e^1 - v_w^1}{h} + \frac{v_n^2 - v_s^2}{h}, \tag{5.11}$$

where h denotes the grid size, which is always taken to be 1 in image processing. Next, we generate further approximations at the midway points, where image information is not directly available. Take the midpoint e, for example:

$$v_e^1 = \frac{1}{|\nabla u_e|}\left[\frac{\partial u}{\partial x}\right]_e \simeq \frac{1}{|\nabla u_e|}\frac{u_E - u_O}{h}, \tag{5.12}$$

$$|\nabla u_e| \simeq \frac{1}{h}\sqrt{(u_E - u_O)^2 + \left[(u_{NE} + u_N - u_S - u_{SE})/4\right]^2}. \tag{5.13}$$

Namely, we approximate $[\partial u/\partial x]_e$ by the central difference scheme, and $[\partial u/\partial y]_e$ by the average of $(u_{NE} - u_{SE})/2h$ and $(u_N - u_S)/2h$. Similar discussion applies to the other three directions N, W, and S.

Therefore, at a pixel O, (5.4) is discretized to

$$0 = \sum_{P \in \Lambda_O} \frac{1}{|\nabla u_p|}(u_O - u_P) + \lambda_e(O)\left(u_O - u_O^0\right), \tag{5.14}$$

where, for example, if $P = E$, then p denotes e. Define

$$w_P = \frac{1}{|\nabla u_p|}, \qquad P \in \Lambda_O, \tag{5.15}$$

$$h_{OP} = \frac{w_P}{\sum_{Q \in \Lambda_O} w_Q + \lambda_e(O)}, \tag{5.16}$$

$$h_{OO} = \frac{\lambda_e(O)}{\sum_{Q \in \Lambda_O} w_Q + \lambda_e(O)}. \tag{5.17}$$

Then (5.14) becomes

$$u_O = \sum_{P \in \Lambda_O} h_{OP} u_P + h_{OO} u_O^0, \tag{5.18}$$

with

$$\sum_{P \in \Lambda_O} h_{OP} + h_{OO} = 1.$$

Equation (5.18) is in the form of a *low pass* filter, which is of course a system of nonlinear equations since the filter coefficients all depend on u.

Freezing the filter coefficients (to linearize the equations), and adopting the Gauss–Jacobi iteration scheme for linear systems, at each step n, we update $u^{(n-1)}$ to $u^{(n)}$ by

$$u_O^{(n)} = \sum_{P \in \Lambda_O} h_{OP}^{(n-1)} u_P^{(n-1)} + h_{OO}^{(n-1)} u_O^{(n-1)}, \tag{5.19}$$

where $h^{(n-1)} = h(u^{(n-1)})$. Since h is a low pass filter, the iterative algorithm is stable and satisfies the *maximum principle* [9]. In particular, the gray value interval $[0, 1]$ is always preserved during the iterating process.

Useful variations of the algorithm can be obtained by altering the definition w_P or $|\nabla u_p|$ in (5.15). For instance, instead of (5.13), we can also try

$$|\nabla u_e| \simeq \frac{1}{h} \sqrt{(u_E - u_O)^2 + [(u_{NE} - u_{SE})/2]^2}.$$

Experiments show that such variations sometimes work better for inpainting sharp edges in the digital setting.

In implementation, as in (5.6), the weights w_P are "lifted" to

$$w_P = \frac{1}{|\nabla u_p|_a} = \frac{1}{\sqrt{a^2 + |\nabla u_p|^2}} \tag{5.20}$$

for some small number a, to avoid a zero divisor in smooth regions. Notice that choosing a large a brings the TV model closer to the harmonic inpainting (especially computationally, since the spatial step size h is set to 1, and u takes values from the finite gray-scale interval $[0, 1]$). In addition, as a gets bigger, the convergence of the iteration scheme speeds up.

The size of the extension domain E is also easily determined. If the image is clean, E can simply be the boundary of the inpainting domain D. Otherwise, to clean up statistical noise and extract reliable image information near the boundary, one can choose E with a reasonable size, e.g., several pixels wide, as practiced in image processing [21]. If, as for the inpainting of an old photo, the entire image is contaminated by noise, then one should take E to be the complement of D, to simultaneously clean and inpaint the photo.

6. Segmentation-based inpainting. The key to image inpainting is the right model for image functions. Image models play a universally crucial role for image restoration problems, such as image denoising, deblurring, and segmentation. In terms of the Bayesian methodology, this is the significance of figuring out an appropriate *prior* model. The link between the Bayesian approach and the variational method is clearly explained in Mumford [32].

In the previous section, the inpainting model has been constructed based on the total variation norm. The main merits of the total variation prior model are its permission of edges and its convenient numerical PDE implementation. In this section, we briefly discuss an inpainting model that is based on Mumford and Shah's [34] object-boundary image model.

An image is considered as the union of a collection of 2-D smooth objects, which meet each other along their edges. Thus in the variational formulation, the regularity functional is no longer in the simple form of

$$R[u] = \int_\Omega r(|\nabla u|) dx$$

as in (5.1). Instead, it imposes the regularity condition on both the edge curves and individual objects:

$$R_{\text{seg}}[u] = \int_{\Omega\setminus\Gamma} r(|\nabla u|, |\Delta u|)dxdy + \mu\text{length}(\Gamma). \tag{6.1}$$

For example, in the Mumford–Shah segmentation model, $r(s,t)$ is taken to be $s^2/2$. (Here we have replaced the Hausdorff measure of Γ by the length, for simplicity.) Thus the unconstrained energy for the segmentation-based image inpainting is

$$J_{\lambda,\mu}[u,\Gamma] = \int_{\Omega\setminus\Gamma} r(|\nabla u|, |\Delta u|)dxdy + \mu\text{length}(\Gamma) + \frac{\lambda}{2}\int_{\Omega\setminus D}(u-u^0)^2 dxdy, \tag{6.2}$$

where D is the inpainting domain. Apparently, it also upholds the three principles.

This segmentation-based inpainting model is a free-boundary problem. Its algorithm and numerical implementation are much more involved than the TV inpainting model. Among the many existing computational methods, recent progress has been made by Chan and Vese [12] based on the level-set method of Osher and Sethian [37].

We now point out the close connection between the TV inpainting model and the segmentation-based inpainting model. In fact, for images which are nearly cartoons, i.e., $|\nabla u_i|$ is negligibly small on each Ω_i, the two models employ almost the same mechanism. Consider the regularity functionals on a test image that represents a black (u_0) disk with radius r_0 in a white background (u_1). Then the TV regularity is

$$\text{TV}[u] = \int_\Omega |\nabla u|dxdy = 2\pi\int_0^\infty |u_r|rdr = 2\pi r_0(u_1-u_0),$$

where we have used the polar coordinates (r,θ). Similarly, the segmentation regularity (for the perfect segmentation) is

$$R_{\text{seg}}[u] = \int_{\Omega\setminus\Gamma} r(|\nabla u|, |\Delta u|)dxdy + \mu\text{length}(\Gamma) = \mu 2\pi r_0.$$

Up to a multiplicative constant (i.e., the edge contrast), the two measures are equivalent. This equivalence holds even for more complex and general image topology as long as the image remains nearly a cartoon. But in terms of numerical implementation, the TV inpainting model is much easier and faster.

7. The connectivity principle and CDD inpainting. Both TV inpainting and segmentation-based inpainting share one drawback. That is, they both fail to realize the so-called *connectivity principle* of the human disocclusion process [11]. See Figure 7.1 for a typical case.

The example in the figure easily explains why the TV and segmentation-based inpainting models fail to realize the connectivity principle when the inpainting scale becomes large. Let u_{dis} and u_{con} denote the disconnected and connected inpainting reconstructions as in the figure. Suppose that $l > w$. Then the TV model prefers u_{dis} to u_{con}, since

$$\text{TV}[u_{\text{con}}] - \text{TV}[u_{\text{dis}}] = 2l - 2w = 2(l-w) > 0,$$

assuming that the black bar has $u_0 = 0$ and the white background $u_1 = 1$. In the same fashion, under the segmentation regularity, we have

$$R_{\text{seg}}[u_{\text{con}}, \Gamma_{\text{con}}] - R_{\text{seg}}[u_{\text{dis}}, \Gamma_{\text{dis}}] = \mu(2l-2w) = 2\mu(l-w) > 0.$$

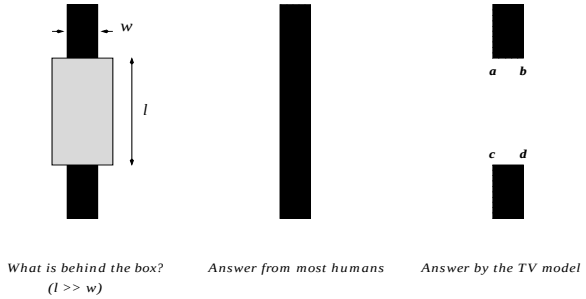

FIG. 7.1. *When $l > w$, the TV and segmentation-based inpaintings both act against the* connectivity principle *of human perception—human observers mostly prefer to have the two disjoint parts connected, even when they are far apart* [24, 36].

Thus the segmentation-based inpainting also biases against the connection.

To overcome such a drawback, Chan and Shen [11] recently proposed a new PDE model based on *curvature driven diffusions* (CDD), which is closely inspired by the TV inpainting model (5.5). The CDD inpainting model is governed by the following PDE:

$$\frac{\partial u}{\partial t} = \nabla \cdot \left[\frac{G(\kappa, x)}{|\nabla u|} \nabla u \right] + \lambda_e (u^0 - u), \quad x \in \Omega, \tag{7.1}$$

where κ is the scalar curvature $\nabla \cdot [\nabla u / |\nabla u|]$. The new ingredient of the CDD model, compared with the TV inpainting model, is the diffusion coefficient $G(\kappa, x)$ which is given by

$$G(\kappa, x) = \begin{cases} 1, & x \in \Omega \backslash D, \\ g(|\kappa|), & x \in D. \end{cases}$$

The choice of a coefficient value of 1 outside the inpainting domain indicates that the model carries out the regular TV denoising task outside D. Meanwhile, $g(s)$ can be any appropriate function that penalizes large curvatures and stabilizes small curvatures inside the inpainting domain. In Chan and Shen [11], it is argued that $g(s)$ must satisfy

$$g(0) = 0, \qquad g(+\infty) = +\infty.$$

Thus, for example, one can choose $g(s) = s^\alpha$ for some $\alpha \geq 1$. Under this condition, the model stretches out bent level lines inside the inpainting domain, outputs connected objects, and therefore realizes the *connectivity principle* (see Figure 10.6, for example).

8. Digital zoom-in based on TV inpainting. Digital zoom-in has wide applications in digital photography, image superresolution, data compression, etc. Zoom-out is a process of losing details or, in the framework of wavelets and multiresolution analysis, a process of projections from fine scales to coarser ones [14, 44]. Zoom-in, on the other hand, is the inverse problem of zoom-out and thus belongs to the general category of image restoration problems. The literature on zoom-ins in image processing has been growing.

One level of zoom-in from a given digital image u^0 of size n by m is to reconstruct a new digital image u of size $2n$ by $2m$ (2 is typical but not unique), so that u^0 can

be the one level zoom-out of u. Thus it is important to know the exact form of the zoom-out operator. Typically, the zoom-out operator consists of two steps: a low pass filtering (or local smooth averaging) of the fine scale image u, followed by a subsampling process leading to the zoom-out u^0 on a coarser grid, a scenario much less strange in wavelet theory [44]. In what follows, we shall assume a direct subsampling zoom-out. That is, the filter is a Dirac δ, and thus the zoom-out is simply a restriction from a $2n$ by $2m$ grid to its n by m double-spaced subgrid.

In contrast to its utility for inpaintings on block domains, continuous modeling becomes less appropriate for the digital setting of zoom-ins. A similar problem has been addressed by Chan, Osher, and Shen [9] for image denoising and enhancement, where a self-contained digital theory for TV denoising was developed and studied. Here we follow the same framework to construct a zoom-in model, which is exactly the digital version of the continuous TV inpainting model.

Let Ω denote the fine grid on which the zoom-in u is to be defined. The grid for the given coarse scale image u^0 is denoted by Ω_0, which is a subgrid of Ω. As in the practice of Markov random fields [3], assign a neighborhood system to Ω, so that each pixel $\alpha \in \Omega$ has its neighborhood N_α, a collection of "nearby" pixels (excluding α itself). For example, we can assign a *rectangular* neighborhood system so that if $\alpha = (i, j)$, then N_α consists of the four pixels $(i, j \pm 1), (i \pm 1, j)$.

At each pixel α, define the local variation as

$$|\nabla_\alpha u| = \sqrt{\sum_{\beta \in N_\alpha} (u_\beta - u_\alpha)^2}.$$

Also define the extended Lagrange multiplier λ_e as a function on the fine grid Ω:

$$\lambda_e(\alpha) = \begin{cases} \lambda, & \alpha \in \Omega_0, \\ 0, & \text{otherwise.} \end{cases}$$

Then the digital TV zoom-in model attempts to minimize the digital energy J_λ over all possible fine scale images u:

$$J_\lambda[u] = \sum_{\alpha \in \Omega} |\nabla_\alpha u| + \sum_{\alpha \in \Omega} \frac{\lambda_e(\alpha)}{2} (u_\alpha - u^0_\alpha)^2. \tag{8.1}$$

For the purpose of comparison, one may also try the digital *harmonic* zoom-in model:

$$J^h_\lambda[u] = \sum_{\alpha \in \Omega} \frac{1}{2} |\nabla_\alpha u|^2 + \sum_{\alpha \in \Omega} \frac{\lambda_e(\alpha)}{2} (u_\alpha - u^0_\alpha)^2. \tag{8.2}$$

As established in [9], the minimization of the digital TV zoom-in energy can be carried out by repeatedly applying the so-called digital TV filter $u \to v = F(u)$: at each pixel α,

$$v_\alpha = F_\alpha(u) = \sum_{\beta \in N_\alpha} h_{\alpha\beta}(u) u_\beta + h_{\alpha\alpha}(u) u^0_\alpha,$$

where the exact formulae for the filter coefficients $h_{\alpha\beta}$ depend on the input u and λ_e and are worked out in [9]. Starting with an arbitrary initial guess $u^{(0)}$ for the zoom-in, we improve its quality by iterating the digital TV filter $u^{(n)} = F(u^{(n-1)})$. As n goes to ∞, $u^{(n)}$ converges to the "best" digital zoom-in of u^0.

As we have noticed, the digital TV zoom-in model (8.1) is almost identical to the continuous TV inpainting model (5.3). The reason we prefer the self-contained digital framework lies in the facts that it is independent of the numerical PDE schemes one applies and always permits a solution (since we are working with finite-dimensional data). The technical difficulty with continuous modeling is that existence is not guaranteed, as discussed by Caselles, Morel, and Shert [4]. The most understandable case is when we choose the H^1 regularity, analogous to the digital version (8.2). Then in the noise-free case, the continuous model is equivalent to finding a harmonic function u on a continuous 2-D domain Ω, which interpolates the given data u^0 on a finite set of pixels. But for harmonic extensions, it is a well-known ill-posed problem to impose both the boundary condition and the 0-dimensional interior interpolation constraint.

9. The inpainting approach to edge-based image coding. In this section, we discuss a very interesting new application of the inpainting technique to edge-based image coding and compression.

Ever since Marr and Hildreth [30], edge has played a crucial role in vision and image analysis, from the classical theory of zero crossings to the more recent theory of wavelets. In image coding, for example, the performance of a scheme is very much determined by its reaction to edges. This viewpoint is further supported by mainstream developments in the current wavelet theory for image coding: Donoho's invention of curvelets and beamlets [16], Mallat's bandlets [28], and Cohen et al.'s tree coding scheme [13].

It would be digressing too much if we tried to explore here the vast literature of image coding and compression. Instead, we now introduce the inpainting approach to (lossy) image coding and compression based on the edge information.

The encoding stage consists of three steps:

- (Edge detection E) Apply an edge detector (Canny's, for example) to detect the edge collection E of a given image u^0. E is typically a set of digital pixels or curves, without good geometric regularities. In addition, we also demand that the physical boundary of the entire image domain Ω belong to the edge collection.
- (Edge tube T) Next, fixing a small constant ϵ, we generate the ϵ-neighborhood T of the edge collection, or as we prefer to call it, *an edge tube*. Digitally, T can be a 1- or 2-pixel thickening of E (see Figure 10.9).
- (Encoding) Finally, we encode the addresses of the tube pixels and use a high bit rate to accurately code the gray values on the tube $u^0\big|_T$.

This encoding scheme creates a large area of "empty seas" where the image information has been wiped out, and thus achieves a high compression rate. In the absence of strong textures and small scale features, the edge collection consists of 1-D piecewise smooth curves. Thus as ϵ tends to zero, the area of the tube T goes to zero, which, theoretically, leads to an infinite compression ratio. Inevitably, such a high compression ratio passes the reconstruction challenge to the decoding scheme. Here we employ the digital TV inpainting scheme to "paint" the uncoded missing information.

To decode, we apply the digital TV inpainting model to the tube T and the gray value data $u^0\big|_T$:

$$\min_u \left[\sum_{\alpha\in\Omega} |\nabla_\alpha u| + \sum_{\alpha\in\Omega} \frac{\lambda_T(\alpha)}{2}(u_\alpha - u^0_\alpha)^2 \right], \tag{9.1}$$

where the extended Lagrange multiplier is

$$\lambda_T(\alpha) = \lambda, \quad \alpha \in T; \qquad 0, \quad \alpha \in \Omega \backslash T.$$

Unlike JPEG or JPEG2000, here the decoding is realized by a variational reconstruction, instead of by a direct inverse transform such as the discrete cosine transform or fast wavelets transform.

The TV norm here has its intrinsic significance. Since during the encoding stage we do not demand any regularity condition on the edge collection, E is typically a messy set without good geometric regularities. Thus the TV norm in the decoding process can straighten the wiggled edges and improve their visual quality.

In Figure 10.9 of the next section, we show a typical example of image decoding based on the TV inpainting model (9.1).

10. Applications of inpainting. For all the inpainting examples of this section, the inpainting domains are given to the algorithm and are initially painted with random guesses, for both the iterative filtering algorithm and the time marching scheme.

10.1. Inpainting a noisy step edge and occluded bars. See Figures 10.1 and 10.2. In the first example, a noisy step edge has been inpainted faithfully by the TV inpainting model. For the second, the occluded bars are recovered as expected.

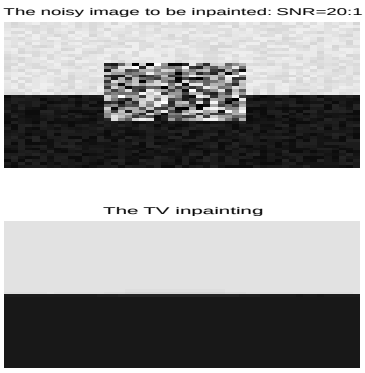

FIG. 10.1. *Inpainting a noisy edge* (10.1).

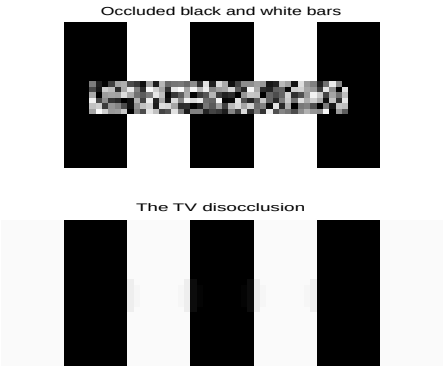

FIG. 10.2. *Inpainting occluded bars* (10.1).

10.2. Inpainting on a topologically complicated domain. See Figure 10.3. This example can also be found in Bertalmio et al. [1]. One can clearly see that the missing circular edges have been linearly approximated by the TV model. The figure easily reminds us the close connection between the TV inpainting model and the problem of minimal surfaces (Giusti [20]).

10.3. Inpainting a noisy scratched photo. See Figure 10.4. The image represents the scanned noisy data of an old scratched photo. As promised, the TV inpainting model can simultaneously denoise the available part of the photo and fill in the missing features. This is the beauty of the TV inpainting: in both the model and algorithm, denoising and inpainting are coherently integrated.

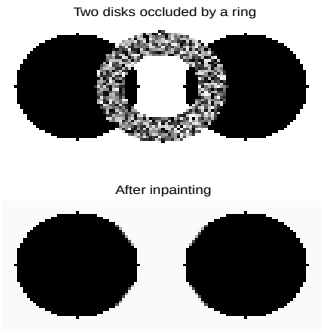

FIG. 10.3. *Inpainting two disks* (10.2).

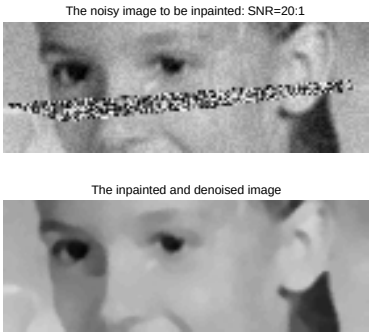

FIG. 10.4. *Inpainting a noisy face* (10.3).

FIG. 10.5. *TV for text removal* (10.4).

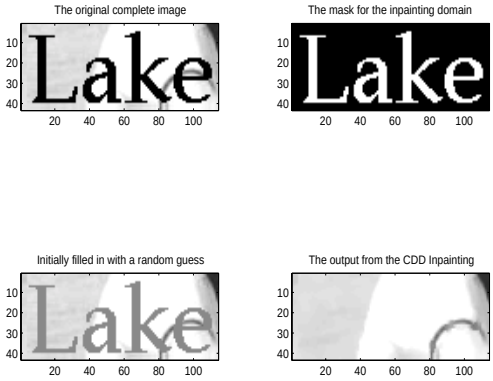

FIG. 10.6. *CDD for text removal* (10.4).

10.4. Removal of thick text. See Figure 10.5. The text string "Lake & Me" has been removed, and the original features occluded by these letters are inpainted. Note that the black rim around the right arm of the T-shirt is not successfully restored by the TV inpainting. The "failure" is due to the scale factor discussed in section 2. The inpainting scale (i.e., the width of a letter in this case) is larger than that of the feature (i.e., the black rim). In Figure 10.6, we have applied the CDD inpainting scheme (section 7) to the same image. For CDD, the connectivity principle is enforced, and therefore the broken rim segments are indeed connected.

10.5. Removal of dense text. See Figure 10.7. The dense text strings have been successfully removed. We feel that this is a very promising application since (a) such problems are typically local due to the small size of the letters, and (b) the number of letters and the complexity of their shapes are well handled by the TV inpainting algorithm since they are easily encoded into the extended Lagrange multiplier λ_e.

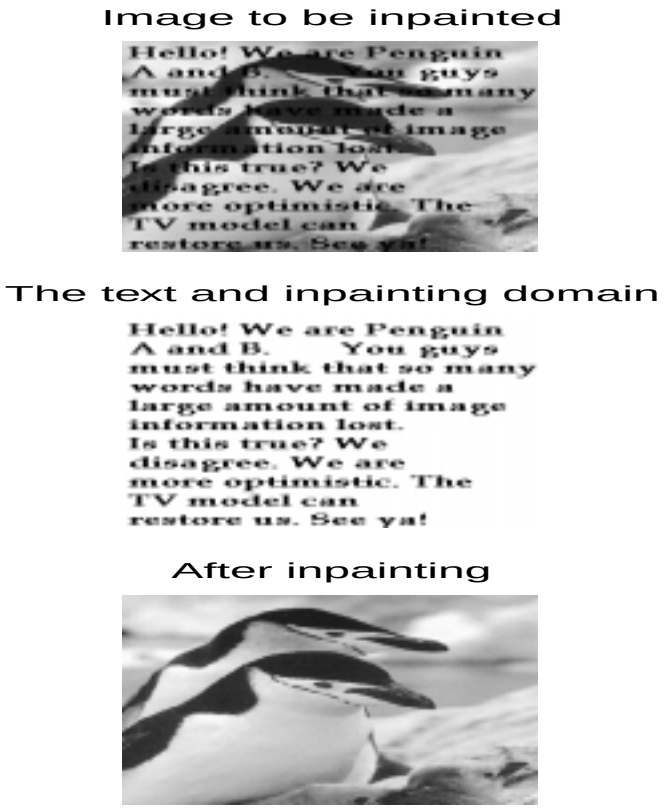

FIG. 10.7. *Removal of dense text* (10.5).

10.6. Digital zoom-in. See Figure 10.8. We apply both the digital TV zoom-in (8.1) and harmonic zoom-in (8.2) to the test image "Lamp" from the image bank of Caltech's Computational Vision Group. It is clear that the TV zoom-in model produces much better visual output in terms of edge sharpness and boundary regularity.

10.7. Edge decoding by inpainting. In Figure 10.9, we show an example of the inpainting approach for image decoding based on edge information. The edge detector we have employed belongs to Canny, which is now a standard MATLAB built-in function. The thickening width described in the previous section is one pixel. This highly lossy coding scheme certainly loses some details of the original image. But remarkably, it faithfully captures the most essential visual information of the image.

Acknowledgments. We owe enormous gratitude to Dr. Marcelo Bertalmio for his generosity in exposing his new work to us through his inspiring talk at UCLA. The present paper would have been absolutely impossible without our initially reading his paper [1]. In addition, the second author would like to thank Professor David Mumford for his encouragement. The authors are also very grateful for inspiration from Professors Stan Osher, Peter Olver, Fadil Santosa, Robert Gulliver, and Dan Kersten during the manuscript revision period.

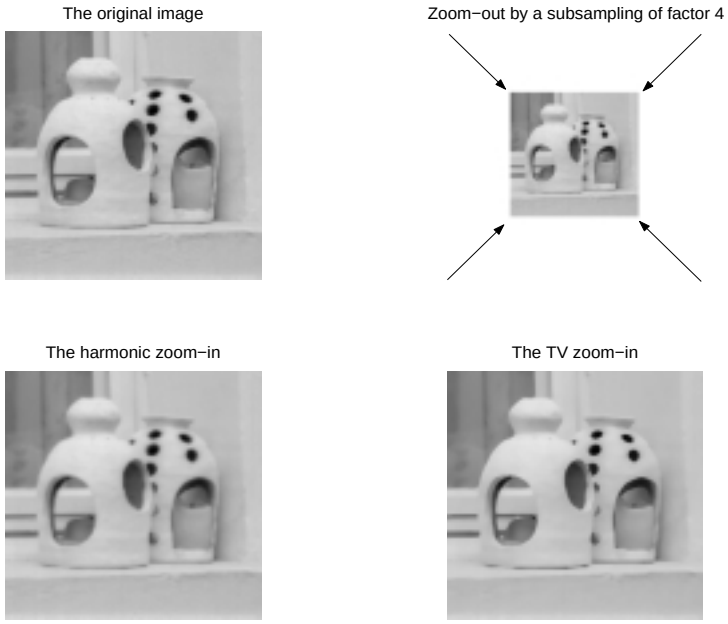

FIG. 10.8. *Digital harmonic zoom-in and TV zoom-in* (10.6).

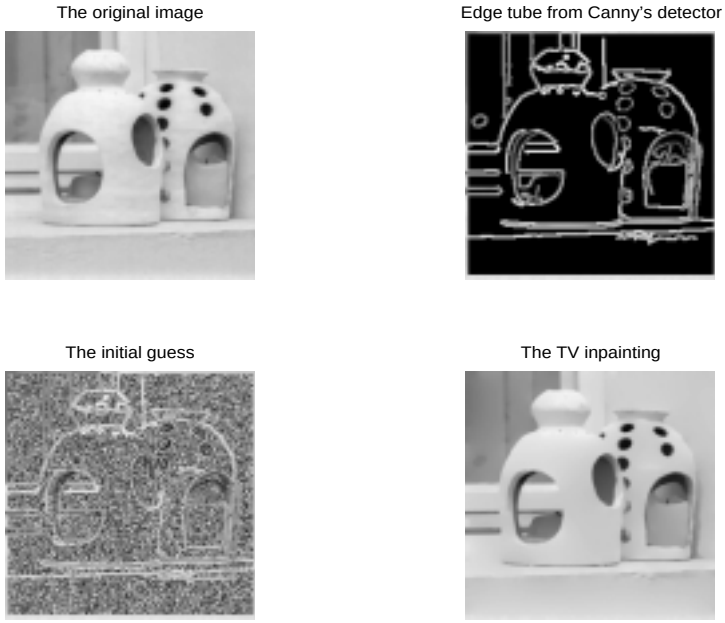

FIG. 10.9. *Edge decoding by TV inpainting* (10.7).

REFERENCES

[1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, *Image inpainting*, in Proceedings of SIGGRAPH 2000, New Orleans, LA, 2000.

[2] P. Blomgren and T. F. Chan, *Modular Solvers for Constrained Image Restoration Problems*, CAM Report 97-52, Department of Mathematics, UCLA, Los Angeles, CA, 1997.

[3] P. Brémaud, *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*, Springer-Verlag, New York, 1998.

[4] V. Caselles, J.-M. Morel, and C. Sbert, *An axiomatic approach to image interpolation*, IEEE Trans. Image Process., 7 (1998), pp. 376–386.

[5] A. Chambolle and P. L. Lions, *Image recovery via total variational minimization and related problems*, Numer. Math., 76 (1997), pp. 167–188.

[6] T. F. Chan, S.-H. Kang, and J. Shen, *Euler's Elastica and Curvature-based Image Inpaintings*, CAM Report 01-12, Department of Mathematics, UCLA, Los Angeles, CA, 2001; available online at www.math.ucla.edu/applied/cam/index.html.

[7] T. F. Chan, S.-H. Kang, and J. Shen, *Total variation denoising and enhancement of color images based on the CB and HSV color models*, J. Visual Comm. Image Rep., 12 (2001).

[8] 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 (1999), pp. 354–367.

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

[10] T. Chan and J. Shen, *Variational restoration of nonflat image features: models and algorithms*, SIAM J. Appl. Math., 61 (2000), pp. 1338–1361.

[11] T. F. Chan and J. Shen, *Non-texture inpainting by curvature driven diffusions (CDD)*, J. Visual Comm. Image Rep., 12 (2001).

[12] T. F. Chan and L. Vese, *A Level-Set Algorithm for Minimizing the Mumford-Shah Functional in Image Processing*, CAM Report 00-13, Department of Mathematics, UCLA, Los Angeles, CA, 2000.

[13] A. Cohen, W. Dahmen, I. Daubechies, and R. DeVore, *Tree Approximation and Optimal Encoding*, Bericht 174, Institut für Geometrie und Praktische Mathematik, Aachen, Germany, 1999.

[14] I. Daubechies, *Ten Lectures on Wavelets,* SIAM, Philadelphia, 1992.

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

[16] D. L. Donoho, *Curvelets*, invited talk at the Workshop on Wavelets, Statistics, and Image Processing, Georgia Institute of Technology, Atlanta, GA, 1999; invited talk at the MSRI workshop on Mathematics of Imaging, Berkeley, CA, 1999. *Beamlets*, invited talk at the IMA Workshop on Image Processing and Low Level Vision, University of Minnesota, Minneapolis, MN, 2000.

[17] G. Emile-Male, *The Restorer's Handbook of Easel Painting*, Van Nostrand Reinhold, New York, 1976.

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

[19] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order,* Springer-Verlag, Berlin, 1977.

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

[21] R. C. Gonzalez and R. E. Woods, *Digital Image Processing,* Addison-Wesley, New York, 1992.

[22] H. Igehy and L. Pereira, *Image replacement through texture synthesis*, in Proceedings of the 1997 IEEE International Conference on Image Processing.

[23] K.-H. Jung, J.-H. Chang, and C. W. Lee, *Error concealment technique using data for block-based image coding*, SPIE, 2308 (1994), pp. 1466–1477.

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

[25] A. C. Kokaram, R. D. Morris, W. J. Fitzgerald, and P. J. W. Rayner, *Detection of missing data in image sequences*, IEEE Trans. Image Process., 11 (1995), pp. 1496–1508.

[26] A. C. Kokaram, R. D. Morris, W. J. Fitzgerald, and P. J. W. Rayner, *Interpolation of missing data in image sequences*, IEEE Trans. Image Process., 11 (1995), pp. 1509–1519.

[27] W. Kwok and H. Sun, *Multidirectional interpolation for spatial error concealment*, IEEE Trans. Consumer Electronics, 39 (1993).

[28] S. Mallat, *Geometrical Image Representations with Bandlets,* invited talk at the IMA Workshop on Image Processing and Low Level Vision, University of Minnesota, Minneapolis, MN, 2000.

[29] A. Marquina and S. Osher, *A new time dependent model based on level set motion for nonlinear deblurring and noise removal*, in Scale-Space Theories in Computer Vision, Lecture Notes in Comput. Sci. 1682, M. Nielsen, P. Johansen, O. F. Olsen, and J. Weickert, eds., 1999, Springer-Verlag, New York, pp. 429–434.
[30] D. Marr and E. Hildreth, *Theory of edge detection*, Proc. Royal Soc. London B, 207 (1980), pp. 187–217.
[31] S. Masnou and J.-M. Morel, *Level-lines based disocclusion*, Proceedings of the 5th IEEE International Conference on Image Processing, Chicago, IL, 1998, pp. 259–263.
[32] D. Mumford, *The Bayesian rationale for energy functionals*, in Geometry Driven Diffusion in Computer Vision, Kluwer Academic, Norwell, MA, 1994, pp. 141–153.
[33] D. Mumford, *Empirical investigations into the statistics of clutter and the mathematical models it leads to*, a lecture for the review of ARO, 1999; also available online at www.dam.brown.edu/people/mumford/research_new.html.
[34] D. Mumford and J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, Comm. Pure Appl. Math., 42 (1989), pp. 577–685.
[35] R. Nevanlinna, *Analytic Functions,* Springer-Verlag, New York, 1970.
[36] M. Nitzberg, D. Mumford, and T. Shiota, *Filtering, Segmentation, and Depth*, Lecture Notes in Comput. Sci. 662, Springer-Verlag, Berlin, 1993.
[37] S. Osher and J. A. Sethian, *Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations*, J. Comput. Phys., 79 (1988).
[38] S. Osher and J. Shen, *Digitized PDE method for data restoration*, in Handbook of Analytic-Computational Methods in Applied Mathematics, Chapman and Hall/CRC Press, Boca Raton, FL, 2000, pp. 751–771.
[39] P. Perona, *Orientation diffusion*, IEEE Trans. Image Process., 7 (1998), pp. 457–467.
[40] L. Rudin and S. Osher, *Total variation based image restoration with free local constraints*, in Proceedings of the 1st IEEE International Conference on Image Processing, Austin, TX, 1994, pp. 31–35.
[41] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.
[42] J. Shen and G. Strang, *The asymptotics of optimal (equiripple) filters*, IEEE Trans. Signal Process., 47 (1999), pp. 1087–1098.
[43] J. Shen, G. Strang, and A. J. Wathen, *The potential theory of several intervals and its applications*, Appl. Math. Optim., 44 (2001), pp. 67–85.
[44] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, MA, 1996.
[45] B. Tang, G. Sapiro, and V. Caselles, *Color Image Enhancement via Chromaticity Diffusion*, Technical report, ECE-University of Minnesota, Minneapolis, MN, 1999.
[46] B. Tang, G. Sapiro, and V. Caselles, *Direction diffusion*, in International Conference in Computer Vision, to appear.
[47] S. Walden, *The Ravished Image*, St. Martin's Press, New York, 1985.
[48] L.-Y. Wei and M. Levoy, *Fast Texture Synthesis Using Tree-Structured Vector Quantization*, Preprint, Computer Science, Stanford University, Stanford, CA, 2000; also in Proceedings of SIGGRAPH 2000.