UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

Mathematical Models for Local Deterministic Inpaintings

Tony Chan
Jianhong Shen

March 2000

CAM Report 00-11

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

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

MATHEMATICAL MODELS FOR LOCAL NON-TEXTURE INPAINTINGS

TONY CHAN AND JIANHONG SHEN *

Dedicated to Stanley Osher

Abstract. Inspired by the recent work of Bertalmio, Sapiro, Caselles, and Ballester [Technical report, ECE-University of Minnesota (1999)] on digital inpaintings, we develop general mathematical models for *local non-texture* inpaintings. Inside smooth regions, inpaintings are connected to the harmonic and bi-harmonic extensions, and inpainting orders are defined and analyzed. For inpaintings involving the recovery of edges, we propose variational models that are closely connected to the *total variation* (TV) restoration model of Rudin, Osher, and Fatemi [*Physica D*, 60 (1992), pp. 259-268] and the Mumford-Shah segmentation model [*Comm. Pure Appl. Math.*, XLII (1989), pp 577-685]. Emphasis is put on the TV inpainting model due to its simplicity in theory and its efficiency in computation and applications. We demonstrate the applications of the inpainting models in restoring scratched old photos, disocclusions in vision analysis, text removal from images, and digital zoomings.

Key words. Inpainting, disocclusion, interpolation, image restoration, variational method, prior model, total variation, segmentation, curvature driven diffusions, digital zoomings, edge coding and image compression.

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

1. Introduction.

Perhaps the best way to explain inpainting is to quote the first paragraph of the very recent paper of Bertalmio, Sapiro, Caselles and Ballester [1] on digital inpaintings:

> "The modification of images in a way that is non-detectable for an observer who does not know the original image is a practice as old as artistic creation itself. Medieval artwork started to be restored as early as the Renaissance, the motives being often as much to bring medieval pictures "up to date" as to fill in any gaps [15, 44]. This practice is called *retouching* or *inpainting*. The objective of inpainting is to reconstitute the missing or damaged portions of the work, in order to make it more legible and to restore its unity [15]."

That is, *inpainting* is to fill in image information on a blank domain D (or several domains), based upon the image information available outside (see Figure 1.1). On such domains, the original painting or picture has been damaged or erased due to aging, scratching or special effects such as objects disappearance. Therefore, inpainting falls in the general category of image restoration, of which, denoising and deblurring are the two more familiar classical examples.

The significance of inpaintings in image processing can be clearly seen from its wide applications in:

(a) digital restoration of ancient paintings for conservation purposes [15, 44];
(b) restoration of old photographs or films with scratches or missing patches [22, 23];
(c) text removal and objects removal in images for special effects [1];
(d) disocclusion in computer vision [27, 32];
(e) digital zooming (Section 9).

* Tony Chan is with Department of Mathematics, UCLA, Los Angeles, CA 90095-1555, *chan@ipam.ucla.edu*; Jianhong Shen is with School of Mathematics, University of Minnesota, Minneapolis, MN 55455, *jhshen@math.umn.edu*. Research supported by grants from NSF under grant number DMS-9973341 and from ONR under N00014-96-1-0277.

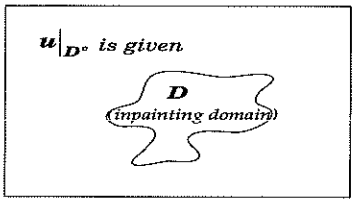


FIG. 1.1. *Inpainting is to paint the missing $u|_D$ on an inpainting domain D based upon what is available on D^c.*

Despite the apparent importance of inpaintings in all these applications, their appropriate mathematical modeling is much less obvious.

On one hand, inpainting is not completely strange to us. It is closely connected to some familiar topics in mathematics, all centered around the task of inferring or restoring the missing information from what is available. Perhaps the two most related examples are:

(a) reconstruction of curves and surfaces based on discrete samples and regularity constraints.

(b) boundary value problems (BVP) which recover functions or distributions in the interior of a domain from the knowledge along its boundary.

On the other hand, inpainting is not simply function interpolation nor a boundary value problem due to the complexity of image functions. For natural images (*clutters*), Mumford proposed to model them by distributions, or "things which can be measured by sensors (i.e. test functions)" [29]. For most texture images, the functions usually contain very rich statistical content, but are far more complicated than the simple samples of some common random fields. In classical linear wavelet analysis, images are believed to be in the space of L^2 (Daubechies [12], Strang and Nguyen [41], for examples), which is also popular in the Fourier method. In the scale-space theory based on anisotropic diffusions, images are considered as in the functional space of bounded variations (Rudin, Osher, and Fatemi [38], and Chambolle and Lions [5]). For most images of man-made objects, piecewise smooth functions are good enough approximations and convenient to work with. Realizing such complexities of image functions helps understand the challenge of the inpainting problem, and the necessity of putting the two restrictive words "local" and "non-texture" in the title of the current paper.

First, this paper shall only discuss inpaintings of non-texture images. That is, we do not intend to work on images with rich texture structures, nor on natural images (clutters), since the inpainting of such images inevitably involves their statistical modeling, which is beyond the scope of the current paper. However, numerical examples in the current paper show that the models and algorithms work well for most real images without sophisticated textures. For texture inpaintings, some recent sample works can be found in Wei and Levoy [45], and Igehy and Pereira [19].

In addition, our models shall only handle *local* inpaintings. Locality, in our opinion, is a crucial concept and we will discuss it in the next section.

The main results in this paper can be summarized into:

(1) For inpaintings of smooth image functions, we define the concept of *inpainting orders*. The introduction of such a simple but necessary concept, can be very useful for any quantitative study on inpaintings, as in numerical analysis and approximation theory.

(2) We establish the basic role of Green's Second Formula for smooth inpaintings, which leads to the *harmonic* and *bi-harmonic* inpainting models. Inpainting orders are rigorously established regardless of the topological complexity of inpainting domains.
(3) We propose three general *principles* for any low-level realistic non-texture inpainting model. They are (a) being local; (b) the capability of restoring narrow broken edges; and (c) robustness to noise.
(4) Based on these general inpainting principles, we propose the *total variation* (TV) inpainting model, which is very similar to the classical restoration model of Rudin, Osher and Fatemi [38] in both theory and algorithm. The model makes it possible to directly inpaint noisy images. It is closely connected to the one proposed by Masnou and Morel [27] in the context of disocclusion, and is simpler and much easier to be implemented digitally, and robust to noise. The variational methodology is distinct from Bertalmio et al.'s approach based on the mechanism of transportation of information from the given data into the inpainting domain [1]. We also present the digital implementation of the TV inpainting model based on numerical PDE's. The linearization technique and a Gauss-Jacobi type of iterative scheme lead to a stable lowpass filtering process, which is typically faster and more stable than schemes based on direct time marchings, and applies easily to digital domains with involved shapes.
(5) We discuss a natural modification of the Mumford-Shah [30] segmentation model for a segmentation based inpainting scheme. We also discuss its link to the TV inpainting model, their common drawback, and the recent work by Chan and Shen [8] on the CDD inpainting scheme, which is designed to overcome such drawback.
(6) We make the connection between digital inpaintings and *digital zoom-in's*. The variational methodology for inpaintings developed in the paper provides a natrual framework for direct digital zoom-in's in the pixel domain.
(7) We develop a new innovative application of inpainting for image compression based on edge coding.

The paper is organized as follows. Section 2 clarifies the meaning of locality and its significance. Connections to computer vision are also made through two familiar classes of examples. Section 3 introduces the *Helmholtz postulation* in computer vision. It is the philosophical bedrock for the variational approach. In Section 4, we study the models and accuracy analysis for inpainting smooth images. Our starting point is Green's Second Formula, which leads to linear and cubic schemes realized by the *harmonic* or *bi-harmonic* inpaintings. In section 5, 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 6, which extends the classical TV restoration model of Rudin, Osher and Fatemi [38]. The digital implementation of the TV inpainting model is also presented. In section 7, we propose a segmentation based inpainting scheme, as inspired by the celebrated segmentation model of Mumford and Shah [30]. Section 8 discusses the last inpainting principle, and briefly introduces the recent work of Chan and Shen [8] on inpaintings based on *curvature driven diffusions* (CDD). In Section 9, we make the link between *digital zoom-in's* and digital 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 [7]. Section 10 explains the new important application of the inpainting technique for edge decoding and image compression. The

last section demonstrates the computational results in restoring scratched old photos, (local) disocclusion in computer vision, texts removal from images, digital zoom-in's and edge decoding.

2. Why Local Inpaintings?. The word "locality" in this paper has two layers of meaning. It implies that both *the class of inpainting problems* we are targeting at and *the class of inpainting schemes* we intend to develop are *local.* We now clarify them from both the viewpoint of numerical mathematics and that of vision analysis.

First, in terms of numerical mathematics, locality is necessary for a faithful reconstruction of the missing image information. The situation is well explained by *Shannon's Sampling Theorem* [33]: *if one expects an accurate reconstruction of a (band-limited) signal, then the sampling distance has to be small enough (according to the band-width).* The same rule applies to inpainting problems. If an inpainting domain is too large, generally (except for the case when the image contains only low frequencies and thus is very smooth and flat), we have already lost some visually meaningful "high-frequency" contents (or small scale features). As a result, any given inpainting scheme can only output a *possible* painting, instead of a visually meaningful approximation to the original image. Since inpainting qualities are mostly judged by human eyes, sometimes a *possible* painting (which is not guaranteed to be close to the original whole image) makes little sense. Inpainting a normal human face with the left (or right) eye region as the inpainting domain is such a good example. Any local inpainting scheme (see the discussion below) can only output a horrible human face with the missing eye region filled in by the surrounding facial color. Thus in this paper, we shall only consider inpaintings whose inpainting domains have small sizes compared to the scales of the missing features.

Secondly, locality in inpaintings is also closely connected to two problems in vision analysis: *local inference* and *the factor of scale.* Similar discussion can be found in the literature of computer vision and human vision (see for example, the psychologist Kanizsa's artistic book [21]).

Local Inference. The locality condition means that our models do not rely on global feature or pattern recognition. Inpainting is thus based on local (with respect to the inpainting domain) information only. For non-texture 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 nearly mirror symmetric along the mouth-nose central line. Though apparent to human visual perception, such patterns are much more difficult and expensive to be caught by digital intelligence, due to their rich variations in scales and structures.

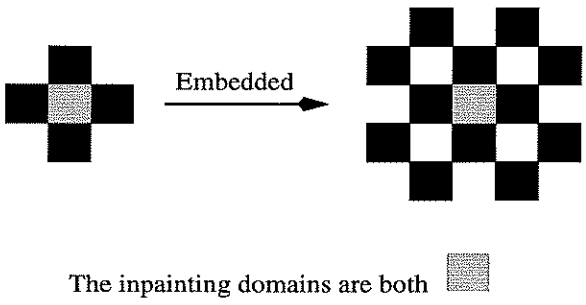


FIG. 2.1. *A local inpainting scheme does not require global pattern recognition.*

Here is a classical example in vision analysis to clarify what we just discussed (refer to Figure 2.1). In the left image to be inpainted, the inpainting (or occluded) domain is the square at the center. To humans, we immediately "see" a black cross despite the missing center, and thus fill in the black color. Most of us agree that this is the best guess. In the right image, we embed the left image into a larger structure. This time, we easily recognize the check-board pattern, and thus fill in the white color. It is called the *completion* of spatial symmetry. The example discussed earlier on inpainting a human face belongs very much to the same category. Without the capability of recognizing global patterns, such outputs cannot be expected.

The complexity of human visual perception parallels that of inpaintings. Any high-level inpainting scheme must be able to carry out pattern recognition. In this paper, the inpainting models should be considered as a low-level one — the inpainting results are independent of global patterns. Therefore, even for the latter case with embedding, the output from the inpainting models shall still be a black cross.

The Factor of Scale. Scale plays a universally significant role in vision analysis and image processing. So it does in the problem of inpainting.

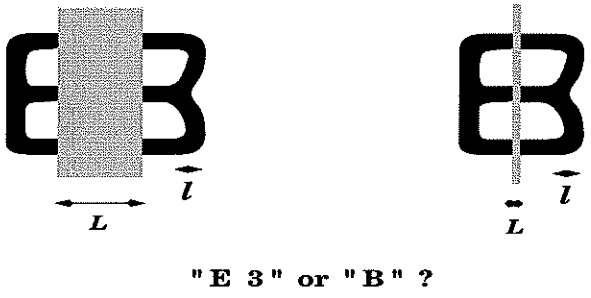


FIG. 2.2. *The effect of the inpainting scale L*

Consider Figure 2.2. In the left image, 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 separate letters "E 3". In the right image, on the other hand, the inpainting scale L is smaller than (or at least comparable with) l. Accordingly, we are more likely to believe that the figure is a broken letter "B." In this example, the non-uniqueness is not caused by global patterns, but by our guess on the correlation among features left there. The controlling parameter is the concept of *scale*. The TV inpainting model and the segmentation based inpainting model developed later output the same results as our human vision does. (However, as discussed in Section 8, sometimes the *Connectivity Principle* must be and can be enforced regardless of the scale factor.)

3. Helmholtz's Postulate in Vision Analysis. In vision research, the following postulate belonging to Helmholtz plays a fundamental role:

The Helmholtz Postulate. *What we perceive is our best guess as to the state of the world given the sensory data.*

This "best guess" principle, in terms of statistics, is the familiar Bayesian view of

the world (see, for example, the well-known paper by the Geman brothers [16]). On the other hand, in terms of the deterministic methodology, such "best guesses" are realized by optimizing energy (or cost) functionals [28]. The difficulty is to come up with perceptually meaningful energy functionals. For image segmentation, the Mumford-Shah [30] functional has proven to be a very successful one. For image restoration involving denoising and deblurring, the Rudin-Osher-Fatemi [38] TV functional is another successful example (flat [37, 38] or non-flat features [9, 36, 42, 43]). For the edge completion from T-junctions in automatic disocclusion, Nitzberg, Mumford, and Shiota [32] recommended Euler's elastica functional.

In this paper, the major inpainting models have been closely inspired by the above existing results on segmentation and restoration, and follow faithfully the Helmholtz Postulate. This is the main difference between our work and Bertalmio et al.'s [1], which is based on the transportation mechanism using a third order differential equation. In the context of computer vision, the mechanism of variational disocclusion first appeared in [32], and has been recently extended by Masnou and Morel [27].

4. Smooth Inpaintings and Green's Second Formula. To develop a rigorous mathematical framework for inpaintings, we start from a simple setting, in which the accuracy of inpainting can be studied. This is the case when the target image function is smooth, or the inpainting domain is contained in the interior of a smooth 2-D object. Besides, studying this simple case can also shed light on the more realistic case coming later.

Let u^0 be a smooth image function defined on a 2-D domain Ω (a rectangular domain, typically). Denote by D the domain to be inpainted and d its diameter. Write the restriction of u^0 on D by $u^0\big|_D$. Then inpainting is to find a function u_D defined on D such that u_D is a good approximation (or *"best" guess*) to $u^0\big|_D$.

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

$$\|u_D - u^0\big|_D\|_\infty = O(d^2). \tag{4.1}$$

Similarly, an inpainting scheme is said to be of k-th order if

$$\|u_D - u^0\big|_D\|_\infty = O(d^{k+1}). \tag{4.2}$$

In this section, we first study such smooth inpainting models.

4.1. Smooth inpainting via Green's Second Formula. Recall that for the 1-dimensional case, harmonic functions on an interval have to be linear functions. Therefore, 1-D linear inpaintings may be carried out equivalently by harmonic extensions. This provides the key for 2-dimensional smooth inpaintings. Here we propose to apply the tool of Green's Second Formula.

Let Δ denote the Laplacian

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

Green's Second Formula on D is

$$\int_D (u\Delta v - v\Delta u)dxdy = \int_\Gamma (u\frac{\partial v}{\partial \mathbf{n}} - v\frac{\partial u}{\partial \mathbf{n}})ds, \tag{4.3}$$

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 $z_0 = (x_0, y_0) \in D$, $G(z_0, z)$, as a function of the "field" point $z = (x, y) \in D$, 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

$$u^0(z_0) = \int_{\Gamma} u^0(z(s)) \frac{\partial(-G(z_0, z))}{\partial \mathbf{n}} ds + \int_D G(z_0, z) \left(-\Delta u^0(z)\right) dz, \tag{4.4}$$

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

In Eq. (4.4), 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 [31]).

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

$$\Delta u^a(z) = \Delta u^0(z), \qquad z \in D; \qquad \text{and} \quad u^a\Big|_{\Gamma} = 0. \tag{4.5}$$

Numerically, 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 standard, but they are indeed fresh for applications in image processing and computer vision, and standard proofs are therefore put down here for completeness. We believe that the complex potential theory will find its wider applications in digital signal and image processing. (For example, recent applications in digital filter design can be found in [39, 40].)

THEOREM 4.1. *Let d denote the diameter of a domain D and $G(z_0, z)$ the associated Green's function for the Poisson equation. Then*

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

The proof is based upon two simple lemmas.

LEMMA 4.2 (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 $z_0, z \in D_1$,*

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

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,

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

since the grounded Green's function is always non-negative. Moreover, $g(z)$ is harmonic inside D_1 because the logarithm singularities at z_0 are canceled out. Therefore $g(z) \geq 0$ for all $z \in D_1$ due to the *extremum principle* of harmonic functions: the minimum is always achieved along the boundary (Gilbarg and Trudinger [17]). This proves the lemma. □

LEMMA 4.3. *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.) □

Proof of Theorem 4.1. 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 4.2 shows that

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

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

$$G_d(z_0, z) = G_1(\frac{z_0}{d}, \frac{z}{d}), \tag{4.6}$$

where G_1, as in Lemma 4.3, is the Green's function for B_1. (This scaling law is true only for the 2-dimensional case.) Therefore, by Lemma 4.3, for any $z_0 \in D$,

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

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 Eq. (4.4), 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 4.1,

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

This validates the assertion.

(b) **Cubic inpainting via Green's Formula.** To improve the accuracy of inpaintings based on Green's Second Formula, we need inpaint the "detail" component u^a missed by the harmonic inpainting.

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

$$u_D^a(z_0) = \int_D G(z_0,z)(-u_D^\Delta(z))dz, \tag{4.7}$$

or equivalently, by solving the grounded Poisson equation

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

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

$$u_D(z) = u^h(z) + u_D^a(z). \tag{4.8}$$

THEOREM 4.4 (Cubic Inpainting). *If u_D^{Δ} is a linear inpainting of Δu^0 on D, then Eq. (4.7) and (4.8) 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

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

for all $z \in D$. Hence,

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

The proof is then complete by Theorem 4.1. □

Remark 2. In the above cubic inpainting process, if the linear inpainting u_D^{Δ} of $\Delta u^0\big|_D$ is realized by a harmonic inpainting, then the cubic inpainting is in fact a bi-harmonic inpainting. That is, $u_D(z)$ solves the following bi-harmonic boundary value problem:

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

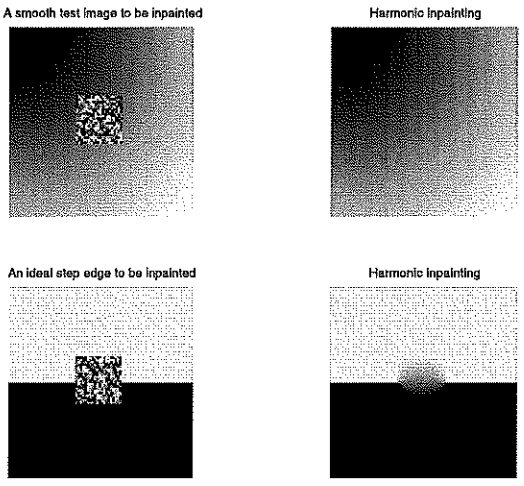


FIG. 5.1. *Harmonic inpaintings of a smooth image ($u = r = \sqrt{x^2 + y^2}$) and an ideal step edge.*

5. 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 not practical since:

(a) Images are *deterministically* not smooth functions. They contain edges and discontinuities (see Figure 5.1).
(b) Images are often *statistically* corrupted by noise.

These properties bring the modeling process back to Helmholtz's "best guess" principle in Section 3. More specifically, a practical inpainting scheme should be able to find the "best" guess of the clean features in the presence of noise, and the "best" guess for the missing features from those in existence.

The practice of best guess also appears in disocclusion and the human inpainting process. In disocclusion, the objects occluded in the scene can only be guessed based on the parts we can see. Similarly, to inpaint ancient paintings, the inpainter has to fill in colors and objects based on the features in existence and a best guess for the missing.

Thus it is natural to build a realistic inpainting model on the basis of the "best guess" principle. As mentioned in Section 3, in the deterministic approach, a best guess is modeled by the optimization of some energy or cost functional.

Faithfully modeling the "best guess" is not so simple since one must deal with high-level vision tasks such as pattern recognition and statistical learning. In the current paper, we shall focus on low-level inpainting models only, which do not rely on high-level vision processings.

With all the above discussion in mind, we summarize three important principles that a low-level (variational) inpainting model must obey. In the coming sections, we shall construct two such models, namely, the total variation (TV) inpainting model, and the segmentation based inpainting model.

(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 for 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 bi-harmonic inpainting are models for *smooth* images. Hence they do not meet Inpainting Principles II and III. They are indeed *local* since only the behavior of u^0 near a neighborhood of Γ is needed (for obtaining the traces of u^0 and Δu^0 on Γ).

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

6. The Total Variation (TV) Inpainting Model.

6.1. Formulation of the TV inpainting model. Let D be an inpainting (open) 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 6.1). Such a setting is motivated by the Inpainting Principles of locality and robustness to noise.

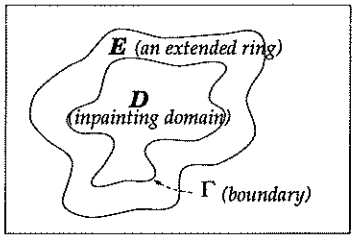


FIG. 6.1. *The TV inpainting model finds the best guess for* $u|_D$ *based on the TV norm on the extended domain* $E \cup D$ *and the noise constraint on* E.

We assume that $u^0\Big|_E$ is contaminated by homogeneous white noise. 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:

$$R[u] := \int_{E \cup D} r(|\nabla u|) dxdy, \tag{6.1}$$

under the fitting (or denoising) constraint on E

$$\frac{1}{\text{Area}(E)} \int_E |u - u^0|^2 dxdy = \sigma^2. \tag{6.2}$$

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 3. 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 [38] if $r(s) = s$.

The variational formulation (6.1) and (6.2) have 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 carefully the regularity functional $R[u]$ or $r(s)$. Along a step edge, ∇u is a 1-dimensional 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 convexity, $\alpha = 1$ is the ideal choice. It leads to the well known total variation (TV) restoration model of Rudin, Osher and Fatemi [38], where $r(s)$ is taken to be s exactly. In this paper, we shall also make the same choice, and call the resulting inpainting model the *TV inpainting model.* It meets all the three inpainting principles.

As in the segmentation model of Mumford and Shah [30] and the TV restoration model of Rudin, Osher and Fatemi [38], it is more convenient to solve the unconstrained TV inpainting problem

$$J_\lambda[u] = \int_{E\cup D} |\nabla u| dxdy + \frac{\lambda}{2}\int_E |u-u^0|^2 dxdy, \tag{6.3}$$

where λ plays a role of the Lagrange multiplier for the constrained variational problem (6.1) and (6.2).

The Euler-Lagrange equation for the energy functional J_λ is

$$-\nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right) + \lambda_e(u-u^0) = 0, \tag{6.4}$$

for all $z=(x,y)\in E\cup D$, plus the Neumann boundary condition [9, 38]. 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

$$\frac{\partial u}{\partial t} = \nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right) + \lambda_e(u^0-u). \tag{6.5}$$

Since λ_e takes two different values, (6.4) or (6.5) is a two-phase problem, and the interface is the boundary Γ of the inpainting domain.

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

$$\nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right) \quad \text{by} \quad \nabla\cdot\left(\frac{\nabla u}{|\nabla u|_a}\right), \tag{6.6}$$

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 (6.1). 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.

From the theoretical point of view, the lifting parameter a also better conditions the TV inpainting model (6.3). In a noise free situation, (6.3) is reduced to a boundary value problem:

$$\nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right) = 0, \quad x\in D; \qquad u\big|_{\partial D} = u^0\big|_{\partial D}. \tag{6.7}$$

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,

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

As a second order equation, its local elliptic symbol σ_a is:

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

where σ_0 is the symbol for the TV model. It is then easy to show that

PROPOSITION 6.1 (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*

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

Therefore, at each pixel away from 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 4. Inspired by Mumford's proposal of Euler's elastica

$$\int_L (\nu + \alpha \kappa^2) ds \tag{6.8}$$

for edge completion in disocclusion (ν and α are constants, and L denotes a permissable edge connection between two detected T-junctions, and κ its curvature), Masnou and Morel in [27] suggested the disocclusion functional

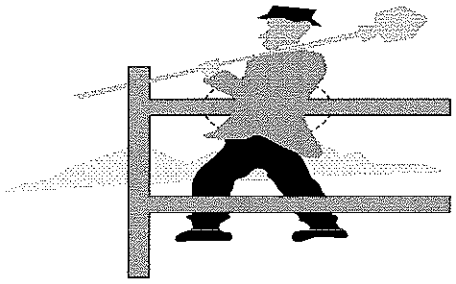
$$\int_\Omega |\nabla v| (1 + |\kappa(v)|^p)) dx dy \qquad \text{with} \qquad v = u \quad \text{outside } D, \tag{6.9}$$

where,
(a) Ω is the entire image domain, and D the occluded region;
(b) $v = v(x, y)$ is a permissable disocclusion and u the given image;
(c) $\kappa(v)$ is the curvature of level lines as given by the first expression in Eq. (6.6).
Like the Laplacian Δv, the curvature $\kappa(v)$ is a second-order feature of v. Therefore, similar to the bi-harmonic inpainting model discussed earlier, the Masnou-Morel model leads to a fourth order Euler-Lagrange equation, and can be considered as a high-order correction of the TV inpainting model (when the image is not contaminated by noise). This fourth-order equation is highly nonlinear and ill-posed, which accounts for Masnou and Morel's preference of the level-line based dynamic programming algorithm over the numerical PDE method.

Remark 5. So far, the TV inpainting model has been solely inspired by the variational formulation for smooth inpaintings and the three inpainting principles. We now further justify the TV norm through a well-known example in human vision phenomena, which may convince us (partially if not completely) that the TV norm does well

approximate our "best guess" in visual perception. This puts the TV model under the general framework of Helmholtz's Postulate in Section 3.

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 of Kanizsa [21]. Its importance for the mathematical understanding and modeling of human vision was first emphasized in Nitzberg, Mumford, and Shiota's systematic work on disocclusion [32]. We have plotted a simple version in Figure 6.2, and name it "Kanizsa's Entangled Man."



Kanizsa's entangled man

FIG. 6.2. *Can the TV inpainting model explain Kanizsa's Entangled Man?*

Figure 6.2 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 et al. [32] 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 stubborn best guess by our visual perception.

The contradiction occurs inside the circled region in Figure 6.2: the fact is that the upper body of the man is behind the fence, while our perception strongly suggests the opposite. Such controversy is apparently caused by the same color shared by the fence and the upper body. So the puzzle is: why does our vision 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, the functional approach based on variational principles thus seems to be more general.

First we simplify the problem to the left image in Figure 6.3. The vertical and horizontal bars model separately the upper body and the fence. Notice the length scales $L > l$, and in Figure 6.2, 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 on the right in Figure 6.3). Now treat D as an inpainting domain and denote by u_D the optimal solution on D obtained

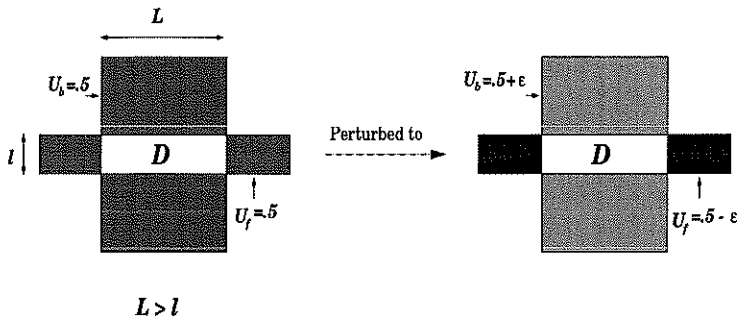


FIG. 6.3. *The model for Kanizsa's Entangled Man.*

from the TV inpainting model with $\lambda = \infty$ (since there is no noise) and E the complement of D. A simple calculation shows that

$$u_D = u_b = 1/2 + \epsilon, \tag{6.10}$$

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

In fact, by the symmetry of the entire domain, it is easy to see that the optimal solution u_D must be a constant, c, say. Then the Maximum Principle [7] requires that $u_f \leq c \leq u_b$. The TV norm (in the distributional sense) of u_D on the closure of D concentrates along the four edges and equals

$$2 \times (|u_f - c| * l + |u_b - c| * L) = [(1 + 2\epsilon)L - (1 - 2\epsilon)l] - (L - l)c. \tag{6.11}$$

We do not care about the TV norm on E because it is a fixed quantity for this noise free inpainting problem. To minimize the TV norm as given in Eq. (6.11), the only choice is $c = u_b = 1/2 + \epsilon$ since $L > l$. This proves the claim.

6.2. Numerical implementation. If the inpainting domain D is empty, then Eq. (6.4) and (6.5) together is exactly the Rudin-Osher-Fatemi [38] denoising and deblurring restoration model. Its theoretical study can be found in Chambolle and Lions [5] and others. Numerical investigations and discussions can be found in [2, 6, 13, 38], and more recent ones in [7, 25, 35]. New applications of the TV model for restoring non-flat image features such as optical flows and chromaticity can be found in the recent papers by Perona [36], Tang, Sapiro, and Caselles [42, 43], and Chan and Shen [9].

In this paper, we have adopted the following numerical scheme for the TV inpainting model (6.4). Here we look for the equilibrium solution directly, instead of by the time marching (6.5), which is usually slow due to the timestep constraints for numerical stability.

As in Figure 6.4, 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

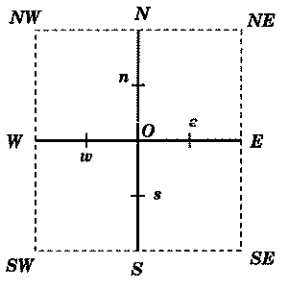


FIG. 6.4. *A target pixel O and its neighbors*

differencing:

$$\nabla \cdot \mathbf{v} = \frac{\partial v^1}{\partial x} + \frac{\partial v^2}{\partial y} \tag{6.12}$$

$$\simeq \frac{v_e^1 - v_w^1}{h} + \frac{v_n^2 - v_s^2}{h}, \tag{6.13}$$

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 mid-point e for example,

$$v_e^1 = \frac{1}{|\nabla u_e|}\left[\frac{\partial u}{\partial x}\right]_e \simeq \frac{1}{|\nabla u_e|}\frac{u_E - u_O}{h}, \tag{6.14}$$

$$|\nabla u_e| \simeq \frac{1}{h}\sqrt{(u_E - u_O)^2 + \left[(u_{NE} + u_N - u_S - u_{SE})/4\right]^2}. \tag{6.15}$$

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 Eq. (6.4) is discretized to

$$0 = \sum_{P \in \Lambda_O} \frac{1}{|\nabla u_p|}(u_O - u_P) + \lambda_e(O)\left(u_O - u_O^0\right), \tag{6.16}$$

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

$$w_P = \frac{1}{|\nabla u_p|}, \qquad P \in \Lambda_O, \tag{6.17}$$

$$h_{OP} = \frac{w_P}{\sum_{P \in \Lambda_O} w_P + \lambda_e(O)}, \tag{6.18}$$

$$h_{OO} = \frac{\lambda_e(O)}{\sum_{P \in \Lambda_O} w_P + \lambda_e(O)}. \tag{6.19}$$

Then Eq. (6.16) becomes

$$u_O = \sum_{P \in \Lambda_O} h_{OP} u_P + h_{OO} u_O^0, \tag{6.20}$$

with

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

Eq. (6.20) is in the form of a *lowpass* filter, which is of course a system of nonlinear equation 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:

$$u_O^{(n)} = \sum_{P \in \Lambda_O} h_{OP}^{(n-1)} u_P^{(n-1)} + h_{OO}^{(n-1)} u_O^{(n-1)}, \tag{6.21}$$

where $h^{(n-1)} = h(u^{(n-1)})$. Since h is a lowpass filter, the iterative algorithm is stable and satisfies the *Maximum Principle* [7]. 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 (6.17). For instance, instead of Eq. (6.15), 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 sharp edges in the digital setting.

Remark 6.

(a) **(Lifting parameter a)** In implementation, as in Eq. (6.6), the weights w_P's are "lifted" to

$$w_P = \frac{1}{|\nabla u_p|_a} = \frac{1}{\sqrt{a^2 + |\nabla u_p|^2}} \tag{6.22}$$

for some small number a (0.01, for example), 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.

(b) **(Lagrange multiplier λ)** For inpaintings of clean images, we can choose λ arbitrarily large. The ideal case is $\lambda = \infty$, which amounts to requiring that

$$h_{OP} \equiv 0, \qquad h_{OO} \equiv 1,$$

and pixels outside the inpainting domain are thus unchanged. To inpaint a noisy image, λ is determined by the noise level, and its choice and efficient estimation have been discussed in [2, 7, 38].

(c) **(Complexity of inpainting domains)** The TV inpainting model and algorithm both easily handle the complexity of the inpainting domain D. Once the inpainting domain is specified by the characteristic function (or *mask*) M

$$M_O = 1, \quad O \in D, \qquad 0, \quad \text{otherwise},$$

then we take λ_e in Eq. (6.4) to be

$$\lambda_e(O) = \lambda(1 - M_O),$$

and the iteration 6.21 goes without being further bothered by the domain complexity.

(d) **(The extension domain E)** The size of the extension domain E is easily determined. If the image is clean, we can simply take E to be the boundary of the inpainting domain D. Otherwise, to clean up the statistical noise, and extract the underlying image information, we need to choose E with a reasonable size, e.g. several pixels wide, as well practiced in image processing [18]. If, as for the inpainting of an old photo, the entire image is contaminated by noise, then we take E to be the complement of D.

7. Segmentation-Based Inpainting. The key to image inpaintings is the right model for image functions. Image models play a universally crucial role in all problems involving image restoration, such as image denoising, deblurring, and segmentation. In terms of the Bayesian worldview, this is the significance of figuring out an appropriate *prior* model. The link between the Baysian statistical methodology and the variational approach is clearly explained in Mumford [28].

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 explore the second important image prior model. This is the *segmentation* prior model as appeared in the celebrated Mumford-Shah segmentation model [30].

In the segmentation prior 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 Eq. (6.1). Instead, it imposes the regularity condition on both the edge curves and individual objects:

$$R_{\text{seg}}[u] = \int_{\Omega\backslash\Gamma} r(|\nabla u|, |\Delta u|)dxdy + \mu\text{length}(\Gamma). \tag{7.1}$$

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

$$J_{\lambda,\mu}[u,\Gamma] = \int_{\Omega\backslash\Gamma} r(|\nabla u|, |\Delta u|)dxdy + \mu\text{length}(\Gamma) + \frac{\lambda}{2}\int_{\Omega\backslash D}(u-u^0)^2dxdy, \tag{7.2}$$

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

The segmentation based inpainting model is a free-boundary problem. Its algorithm and numerical implementation are much more involved than the TV inpainting. Among the many exisiting computational methods, here we outline the level-set method of Osher and Sethian [34], as recently applied to the numerical segmentations by Chan and Vese [10], where interested readers can find more details.

From now on, we assume that we take $r(s,t) = s^2/2$ in the segmentation based inpainting, as in the classical Mumford-Shah segmentation model.

The level-set approach for the segmentation based inpainting relies on the active contour method. Starting with an initially guessed contour $\Gamma(0)$, we let $\Gamma(t)$ evolve to converge to the optimal edge collection Γ. At each time t, $\Gamma(t)$ partitions the whole image domain Ω to disjoint open connected components $\Omega_i(t)$'s so that

$$\Omega\backslash\Gamma(t) = \bigcup_{i=1}^{N(t)} \Omega_i(t),$$

where due to merging and splitting, the number $N(t)$ of connected components also varies with time. Merging and splitting are crucial during the evolution of the initial guess, which makes the level-set method an ideal tool for numerically computing the curve evolution (or *active contour*).

Knowing $\Gamma(t)$ for each time t, we then minimize the inpainting energy $J_{\lambda,\mu}[u,\Gamma(t)]$ on the free variable u, by solving on each connected component $\Omega_i(t)$ the elliptic equation with Nuemann condition:

$$\begin{cases} \Delta u^{(i)} + \lambda_e(x)(u^0 - u^{(i)}) = 0 \\ \left.\frac{\partial u^{(i)}}{\partial \mathbf{n}}\right|_{\partial\Omega_i} = 0. \end{cases} \tag{7.3}$$

Here, as in the TV inpainting model, $\lambda_e(x)$ denotes the extended Lagrange multiplier: $\lambda \times 1_{\Omega\setminus D}$. With $u(t)$ and $\Gamma(t)$, we are then able to evolve $\Gamma(t)$ by the gradient descent method, which is symbolically given by:

$$\frac{d\Gamma}{dt} = -\frac{\partial J_{\lambda,\mu}}{\partial \Gamma}.$$

It allows $\Gamma(t)$ to evolve a numerical time step Δt. Then for the new $\Gamma(t+\Delta t)$, we can repeat the process to further advance the evolution $\Gamma(t)$ and $u(t)$.

In the level-set computational method, of course, $\Gamma(t)$ is always encoded by its level-set function $\phi(x,t)$, and naturally, the above evolution equation for the edge collection $\Gamma(t)$ passes on to $\phi(x,t)$. As well known in the literature, the level-set method eases the merging or splitting of domains and edge curves, thus is suitable for the segmentation based inpainting problem. We refer to [10] for more details on the level-set implementation in the context of image segmentation.

Finally, we 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., on each Ω_i, $|\nabla u_i|$ is negligibly small, the two models almost employ the same mechanism. To get the point, consider the regularity functionals on a test image which represents a black (u_0) disk with radius r_0 in a white background (u_1). Then, the TV regularity is

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

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

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

Up to a multiplicative constant, the two measures are equivalent. This is true 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.

8. The Connectivity Principle and the CDD Inpainting. Both the TV inpainting and segmentation based inpainting share one drawback. That is, they both fail to realize the so called *Connectivity Principle* in the human disocclusion process [8]. See Figure 8.1 for a typical case.

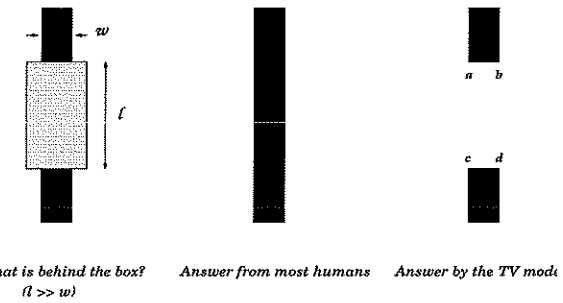


FIG. 8.1. *When $l > w$, the TV and segmentation based inpaintings both act against the* Connectivity Principle *of human perception — humans mostly prefer to have the two disjoint parts connected, even when they are far apart [21, 32].*

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 $l > w$. Then, the TV model perfers 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

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

Thus the segmentation based inpainting also biases against the connection.

To overcome such drawback, Chan and Shen recently proposed a new PDE model based on *curvature driven diffusions* (CDD), which is inspired by the TV inpainting model Eq. (6.5). The CDD inpainting model is governed by the following PDE:

$$\frac{\partial u}{\partial t} = \nabla \cdot \left[\frac{G(\kappa, x)}{|\nabla u|} \nabla u \right] + \lambda_e (u^0 - u), \quad x \in \Omega, \tag{8.1}$$

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 new 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 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 "erases" large curvatures and stablizes small curvatures inside the inpainting domain. In Chan and Shen [8], 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 the condition, the model stretches out bended level lines inside the inpainting domain, and outputs connected objects. Thus the CDD inpainting model realizes the *Connectivity Principle* (see Figure 11.6, for example). More details on the CDD inpainting scheme can be found in [8].

9. Digital Zoom-in and Digital Inpainting. Digital zoom-in has wide applications in digital photography, image superresolution, and 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 [12, 41]. Zoom-in, on the other hand, is the inverse problem of zoom-out, and thus belongs to the general category of image restoration problems.

There has been a growing literature on zoom-in's in image processing. Here we shall only focus on its connection to inpaintings in the current paper.

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 lowpass 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 [41]. 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.

Unlike for inpaintings on domains, continuous modeling becomes less appropriate for the digital setting of zoom-in's. A similar problem has been addressed earlier by Chan, Osher and Shen [7] for image restorations, where a self-contained digital theory for the TV denoising has been developed and studied. Here we follow the digital framework in [7] to construct a zoom-in model based on the digital TV norm, which is formally, as we shall see below, identical to 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 is to minimize the digital energy J_λ over all possible fine scale images u:

$$J_\lambda[u] = \sum_{\alpha \in \Omega} |\nabla_\alpha u| + \sum_{\alpha \in \Omega} \lambda_e(\alpha)(u_\alpha - u^0_\alpha)^2. \tag{9.1}$$

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

$$J^h_\lambda[u] = \sum_{\alpha \in \Omega} \frac{1}{2} |\nabla_\alpha u|^2 + \sum_{\alpha \in \Omega} \lambda_e(\alpha)(u_\alpha - u^0_\alpha)^2. \tag{9.2}$$

As established in [7], 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 coefficents $h_{\alpha\beta}$'s depend on the input u and λ_e, and are worked out in [7]. Starting with an arbitrary initial guess $u^{(0)}$ for the zoom-in, we polish 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. More details regarding the digital TV filter and its algorithm can be found in [7]. A numerical example showing these zoom-in schemes is presented in Figure 11.8 of the next section.

As we have noticed, the digital TV zoom-in model (9.1) is almost identical to the continuous TV inpainting model (6.3). The reason we prefer the self-contained digital framework lies in the facts that it is independent of numerical PDE schemes one applies and always permits a solution (since we are working with finite-dimensional data). The technical difficulty with the continuous modeling is that there may exist no solution to the cost functional, as discussed by Caselles, Morel and Shert [4]. The most understandable case is when we choose the H^1 regularity, analoguous to the digital version (9.2). 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.

10. Inpainting, Edge Coding and Image Compression. In this section, we discuss a very interesting application of the inpainting technique in edge decoding and image compression. We must say that the work presented below is still very preliminary.

Ever since David Marr [26], edge has always been playing 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 better supported by the main stream development in the wavelet theory for image coding: Donoho's invention of curvelets and beamlets [14], Mallat's bandlets (Invited talk at 2000 IMA Workshop on Image and Vision Analysis, University of Minnesota, MN), and Cohen, Dahmen, Daubechies and DeVore's tree coding scheme [11].

It will be digressing too much if we intend 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 (see Figure 11.9). In addition, we also demand the physical boundary (i.e., the four sides of the image domain Ω) to 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 11.9 and 11.10).

– (Encoding) Finally, we encode the addresses of the tube pixels and use high bit rate to accurately code the gray values on the tube $u^0|_T$.

As we see, the 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 high compression ratio brings the challenge to the decoding scheme. Here we employ the digital inpainting scheme to paint the vast empty seas.

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

$$\min_u \left[\sum_{\alpha \in \Omega} |\nabla_\alpha u| + \sum_{\alpha \in \Omega} \frac{\lambda_T(\alpha)}{2} (u_\alpha - u^0_\alpha)^2 \right], \tag{10.1}$$

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, typically E is a messy set without good geometric regularities. Thus the TV norm in the decoding process can straighten the messy edges and improve their visual smoothness.

In Figure 11.9 and 11.10 of the next section, we show two examples of image decoding based on the TV inpainting model 10.1. The test images are again taken from the publicly available web site of Caltech's computational vision group.

11. Applications and Examples. All the examples in this section are generated by the TV inpainting model, the CDD inpainting model, and the digital zoom-in model discussed in the paper. The inpainting domains are given to the algorithm. We have initially painted the missing domains by random guesses for both the filter iteration algorithm and the time marching scheme (of the CDD inpainting).

11.1. Inpainting a noisy step edge and occluded bars. See Figure 11.1 and 11.2. 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.

11.2. Inpainting on a topologically complicated domain. See Figure 11.3. 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. In high order models like Masnou and Morel's for disocclusion [27], a missing circular edge is approximated by an elastica [32], or a similar geodesic curve.

11.3. Inpainting a noisy scratched photo. See Figure 11.4. The image represents the scanned noisy data of an old scratched photo of a human's face. 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.

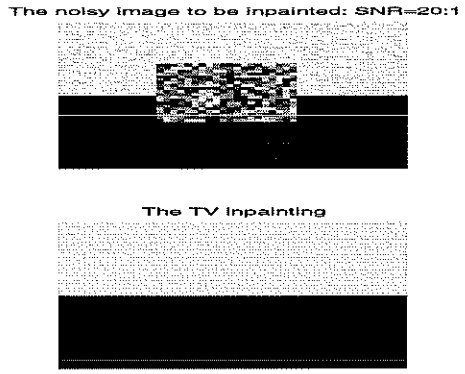


FIG. 11.1. *Inpainting a noisy broken edge (11.1).*

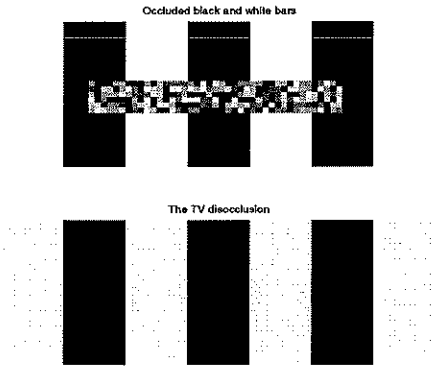


FIG. 11.2. *Inpainting broken bars (11.1).*

11.4. Removal of thick text. See Figure 11.5. The text string "Lake & Me" has been removed and the original occluded features are inpainted. Note that the black rim of the right arm of the T-shirt is not successfully restored by the TV inpainting. The "failure" is connected to the key concept of *scale* 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 11.6, we have applied the CDD inpainting scheme (Section 8) to the same image. For CDD, the Connectivity Principle is enforced and therefore the broken rim segments are indeed connected.

11.5. Removal of dense text. See Figure 11.7. 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; (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.

11.6. Digital zoom-in. See Figure 11.8. We apply both the digital TV zoom-in (9.1) and harmonic zoom-in (9.2) 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 results in terms of edge sharpness and boundary regularity.

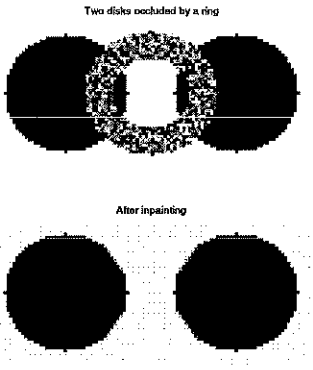


FIG. 11.3. *Inpainting two black disks occluded by a flower ring (11.2).*

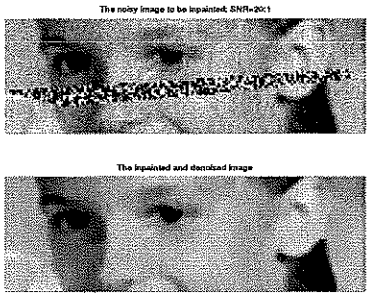


FIG. 11.4. *Inpainting a noisy real image (11.3).*

11.7. Edge decoding by inpainting. In Figure 11.9 and 11.10, we have shown two examples of the inpainting approach for image decoding based on the edge information. The edge detector we have employed belongs to Canny, which is now a standard MATLAB built-in function. The edge thickening width described in the previous section is one pixel. The lossy coding scheme certainly mollifies the original image functions, but most importantly, does catch the essential visual information in the original images.

12. Closing Remarks. We would like to make a few more comments before we end this paper.

(A) The terminology of "inpainting" was first coined by Bertalmio, Sapiro, Caselles, and Ballester [1]. In the literature of image processing, different groups have used different words. Some of the most common ones include *image interpolation* [4, 23], *image replacement* [19], and *disocclusion* [27]. We prefer to use the word "inpainting," not only because it easily reminds the readers of "restoring (ancient) paintings," but also that other terminologies are often either too broad in meaning or of a too high level in terms of vision analysis.

(B) One important application of inpaintings that has recently come to the authors' attention is the problem of *error concealment* in communication systems [20, 24]. In packet based communication networks, the transmission of images or videos is often degraded by the loss of *data blocks*. The problem of recovering these lost data blocks is loosely classified as *error concealment*. One can read [20, 24] for

FIG. 11.5. *Removal of thick text (11.4): The TV inpainting.*

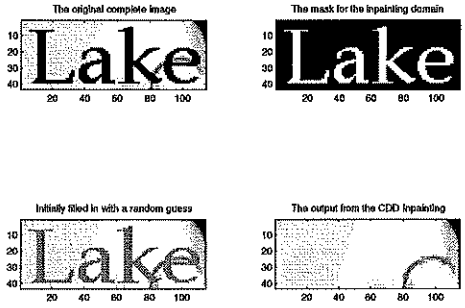


FIG. 11.6. *Removal of thick text (11.4): The CDD inpainting.*

the typical literature of the field. Apparently, like for digital zoom-in (Section 9), the continuous TV inpainting model can lead to a new error concealment scheme by an appropriate digitization.

(C) In terms of level-line connection, both the TV and segmentation based inpainting schemes discussed in the current paper are only of first order. That is, they reconstruct smooth missing edges inside an inpainting domain using straight line segments. Applying Euler's elastica can improve the smoothness (Remark 4 or [27]), but at a much higher computational cost. Numerical PDE schemes are under development by our research group for solving the fourth order Euler-Lagrange equation.

(D) One important research topic to be done is to develop inpainting schemes which combine both texture modeling and the variational/PDE approach. The goal is to inpaint more general images which contain textures.

Acknowledgments. We owe enormous gratitude to Dr. Marcelo Bertalmio for his generosity of exposing his new work to us through his inspiring talk at UCLA. The present paper is absolutely impossible without the inspiration from his paper [1]. The authors are also deeply grateful to the reviewers who offered very generous and constructive suggestions. In addition, the second author also thanks Professor David Mumford for his encouragement.

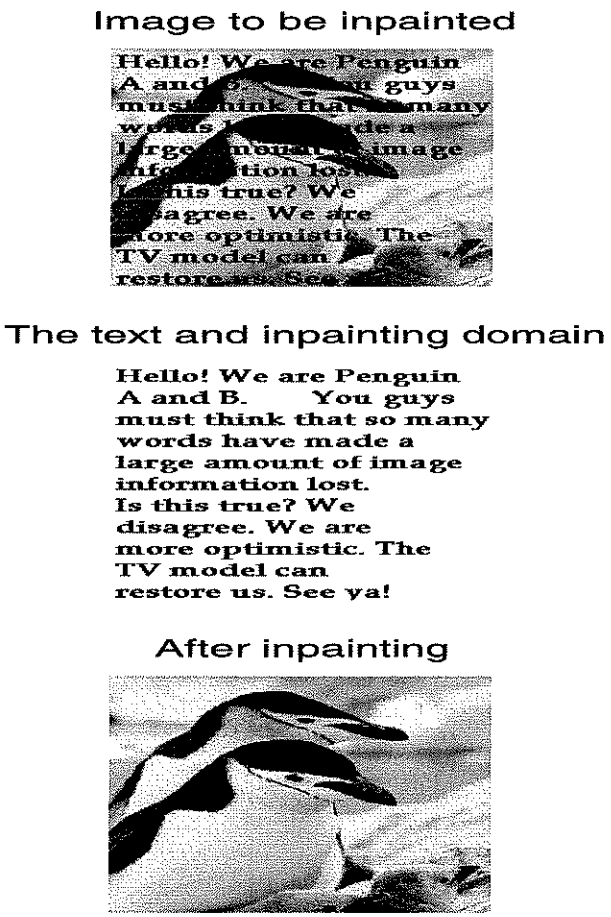


FIG. 11.7. *Removal of dense text (11.5).*

REFERENCES

[1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. Technical report, ECE-University of Minnesota, 1999.

[2] P. Blomgren and T. Chan. Modular solvers for constrained image restoration problems. CAM Report 97-52, Department of Mathematics, UCLA, 1997.

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

[4] V. Caselles, J.-M. Morel, and C. Sbert. An axiomatic approach to image interpolation. *IEEE Trans. Image Processing*, 7(3):376–386, 1998.

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

[6] T. Chan and P. Mulet. On the convergence of the lagged diffusivity fixed point method in image restoration. *SIAM J. Num. Anal.*, 36(2):354–367, 1999.

[7] T. Chan, S. Osher, and J. Shen. The digital TV filter and nonlinear denoising. *IEEE Trans. Image Processing,* in press, 2001.

[8] T. Chan and J. Shen. Non-texture inpainting by curvature driven diffusions (CDD). *J. Visual Comm. Image Rep.,* submitted, 2000.

[9] T. Chan and J. Shen. Variational restoration of non-flat image features: models and algorithms. *SIAM J. Appl. Math.,* in press, 2001.

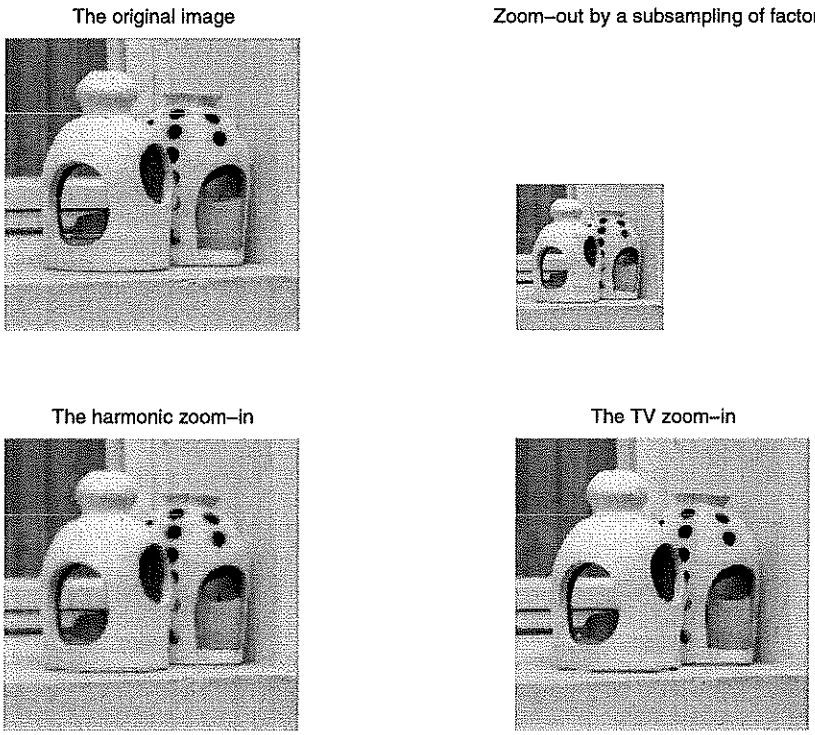


FIG. 11.8. *Digital harmonic zoom-in and TV zoom-in (11.6).*

[10] T. Chan and L. Vese. A level-set algorithm for minimizing the Mumford-Shah functional in image processing. UCLA CAM Report 00-13, Department of Mathematics, 2000.
[11] A. Cohen, W. Dahman, I. Daubechies, and R. DeVore. Tree approximation and optimal encoding. *Institut für Geometrie und Praktische Mathematik,* Bericht Nr. 174, 1999.
[12] I. Daubechies. *Ten Lectures on Wavelets.* SIAM, Philadelphia, 1992.
[13] 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.
[14] D. L. Donoho. Curvelets. Invited talk at GeoTech workshop on *Wavelets, Statistics, and Image Processing*; Invited talk at MSRI workshop on *Mathematics of Imaging*; 1999. Beamlets, Invited talk at IMA workshop on *Image Analysis and Low Level Vision*, 2000.
[15] G. Emile-Male. *The Restorer's Handbook of Easel Painting.* Van Nostrand Reinhold, New York, 1976.
[16] 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.
[17] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order.* Springer-Verlag, Berlin, 1977.
[18] R. C. Gonzalez and R. E. Woods. *Digital Image Processing.* Addison-Wesley, New York, 1992.
[19] H. Igehy and L. Pereira. Image replacement through texture synthesis. *Proceedings of 1997 IEEE Int. Conf. Image Processing.*
[20] K.-H. Jung, J.-H. Chang, and C. W. Lee. Error concealment technique using data for block-based image coding. *SPIE*, 2308:1466–1477, 1994.
[21] G. Kanizsa. *Organization in Vision.* Praeger, New York, 1979.
[22] 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(4):1496–1508, 1995.
[23] 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(4):1509–1519, 1995.
[24] W. Kwok and H. Sun. Multidirectional interpolation for spatial error concealment. *IEEE*

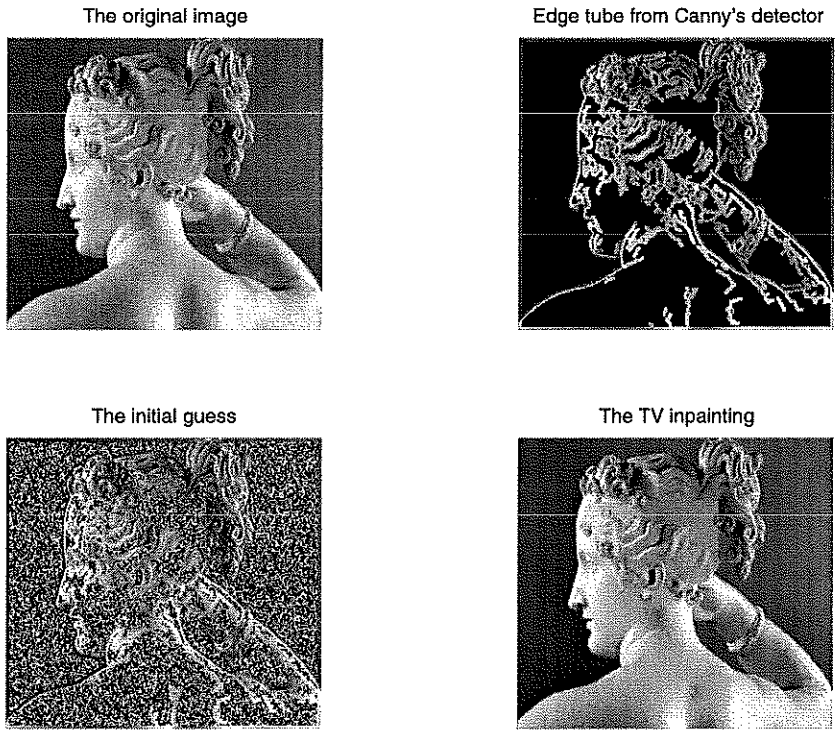


FIG. 11.9. *Edge decoding by TV inpainting: Example 1 (11.7).*

Trans. Consumer Electronics, 39(3), 1993.

[25] A. Marquina and S. Osher. *Lecture Notes in Computer Science*, volume 1682, chapter "A new time dependent model based on level set motion for nonlinear deblurring and noise removal", pages 429–434. 1999.

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

[27] S. Masnou and J.-M. Morel. Level-lines based disocclusion. *Proceedings of 5th IEEE Int'l Conf. on Image Process., Chicago*, 3:259–263, 1998.

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

[29] D. Mumford. Empirical investigations into the statistics of clutter and the mathematical models it leads to. A lecture for the review of ARO, *www.dam.brown.edu/people/mumford/research_new.html*, 1999.

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

[31] R. Nevanlinna. *Analytic Functions*. Springer-Verlag, New York, 1970.

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

[33] A. V. Oppenheim and R. W. Schafer. *Discrete-time signal processing*. Prentice Hall, New Jersey, 1989.

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

[35] S. Osher and J. Shen. *Handbook of Analytic-Computational Methods in Applied Mathematics*, chapter "Digitized PDE method for data restoration", pages 751–771. Chapman & Hall/CRC, FL., 2000.

[36] P. Perona. Orientation diffusion. *IEEE Trans. Image Process.*, 7(3):457–467, 1998.

[37] L. Rudin and S. Osher. Total variation based image restoration with free local constraints.

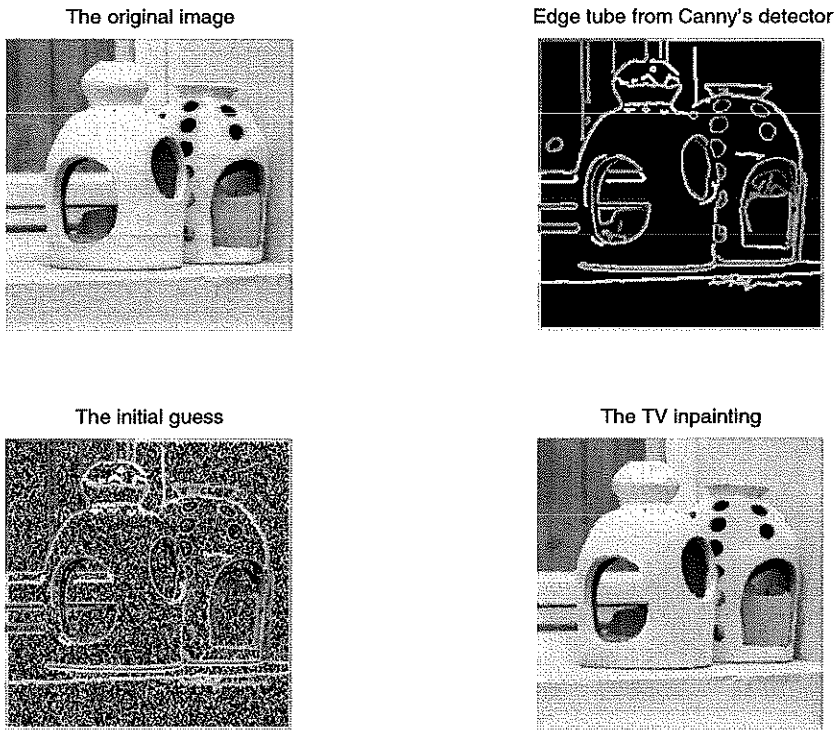


FIG. 11.10. *Edge decoding by TV inpainting: Example 2 (11.7).*

Proc. 1st IEEE ICIP, 1:31–35, 1994.

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

[39] J. Shen and G. Strang. The asymptotics of optimal (equiripple) filters. *IEEE Trans. Sig. Proc.*, 47(4):1087–1098, 1999.

[40] J. Shen, G. Strang, and A. J. Wathen. The potential theory of several intervals and its applications. *Appl. Math. Optim.*, in press, 2001.

[41] G. Strang and T. Nguyen. *Wavelets and Filter Banks.* Wellesley-Cambridge Press, Wellesley, MA, 1996.

[42] B. Tang, G. Sapiro, and V. Caselles. Color image enhancement via chromaticity diffusion. Technical report, ECE-University of Minnesota, 1999.

[43] B. Tang, G. Sapiro, and V. Caselles. Direction diffusion. *International Conference in Computer Vision,* to appear, 1999.

[44] S. Walden. *The Ravished Image.* St. Martin's Press, New York, 1985.

[45] L.-Y. Wei and M. Levoy. Fast texture synthesis using tree-structured vector quantization. Preprint, Computer Science, Stanford University, 2000; Also in *Proceedings of SIGGRAPH 2000.*