from *SIAM News*, Volume 36, Number 5, June 2003

Inpainting and the Fundamental Problem of Image Processing

By Jianhong (Jackie) Shen

"Inpainting" is the art world's term for what researchers in image processing call "image interpolation." In image processing, the fundamental problem is to answer the question "What do we mean by images?" Without knowing what images are, we could not hope to reconstruct missing data—this is the connection between inpainting and the fundamental problem of image processing.

As part of a report on SIAM's first conference on imaging, Guillermo Sapiro (*SIAM News*, Vol. 35, No. 4, May 2002) gave a detailed account of the motivations and numerous applications of image inpainting. It was Sapiro's group, in describing their first third-order PDE-based inpainting model [1], who first borrowed the term "inpainting'' from museum restoration artists. At UCLA, we have taken a different approach, developing inpainting models based on the Bayesian rationale, where the answer to the key question in imaging processing—that is, the identification of appropriate image priors—plays a critical role. Our approach is briefly described here; most of the papers mentioned can be found on the Web site of UCLA's image processing group (www.math.ucla.edu/~imagers).

The Inpainting Problem

Let Ω denote a complete image domain, often a rectangular area on your computer screen, or more generally a finite Lipschitz domain in $\mathbb{R}^2$. Certain factors, such as object occlusion in visual fields and packet loss in wireless communication, result in a subset G of Ω for which image data are missing or inaccessible.

The goal of inpainting is to recover the original ideal image u on the entire domain Ω, based only on the partial (and usually distorted) observation $u_0|_{\Omega\backslash G}$.

Because most objects are not transparent, human observers experience the occlusion effect almost all the time. Still, we see the world as perfectly ordered and integrated, rather than as a cluttered landscape of independent discrete pieces. This is nature's answer to the inpainting problem. Vision and cognitive scientists believe that human beings, without being aware of it, constantly and cleverly apply the rules of Bayesian inferencing and decision making.

Ingredients of Bayesian Inpainting

In the Bayesian framework, inpainting is defined as maximizing the posterior probability $p(u|u_0, G)$. By Bayes's formula,

$$p(u|u_0,G) = \frac{p(u_0|u,G)\,p(u|G)}{p(u_0|G)}.$$

In most applications, the mechanism leading to information loss is independent of the image content, implying that $p(u|G) = p(u)$.

Once u_0 and G are given, $p(u_0|G)$ is simply a normalization constant. In essence, then, we are seeking to maximize the product of the *data model* $p(u_0 \mid u, G)$ and the *image prior model* $p(u)$. A typical data model in image processing involves blurring followed by white noise pollution: $u_0|_{\Omega\backslash G} = (Ku + n)|_{\Omega\backslash G}$. Here K is a blurring operator that is linear and *lowpass*: $K1 = 1$; n is additive Gaussian white noise.

Working with logarithm likelihood functions $E = -1/\beta \ln p$, we want to minimize

$$E[u \mid u_0, G] = E[u_0 \mid u, G] + E[u],$$

up to an additive constant. If $\beta = 1/(kT)$, with k being the Boltzmann constant and T the absolute temperature, then $p = 1/Z \exp(-\beta E)$ is formally the Gibbs formula from statistical mechanics that links the energy of an ensemble to its likelihood. (Here Z is the partition function defining the free energy.)

With Gaussian noise the data model becomes

$$E[u_0|u,G] = \frac{1}{2\sigma^2 |\Omega\backslash G|}\int_{\Omega\backslash G}(u_0 - Ku)^2\,dx,$$

where $dx = dx_1dx_2$ is the Lebesgue area element, σ^2 the variance, and $|\Omega\backslash G|$ the Lebesgue measure. Essentially, the key to the inpainting problem is to employ a suitable image prior $p(u)$ or $E[u]$—which is the fundamental problem in image processing.

Fundamental Problem of Image Processing

What are images, mathematically speaking? Engineers have worked in image processing for nearly a century without asking the question. Why should mathematicians care about it now? In our opinion, the value of asking and answering this fundamental question is no different from that of laying down the Hilbert space foundation for quantum mechanics. Without a satisfactory answer, it may never be possible for image processing to become a genuine new branch of mathematics.

A general problem in image processing can be modeled as an input–output system [5]:

$$U_0 \rightarrow \boxed{\text{Image Processor } \mathcal{T}} \rightarrow U,$$

where the input U_0 could be a single image or an observed image sequence; τ is a linear or nonlinear image processor, such as restoration and compression; and U is the image feature to be determined. In the inpainting case, for example, $U_0 = (u_0, G)$ and $U = u$. Knowing the class of objects U_0 and U (or the definition and range domains of τ) is thus crucial for effective mathematical modeling, analysis, and computation of τ.

Three major approaches have been taken to the fundamental problem:

Physical simulation. Images are generated via simulation of the underlying physical, chemical, or biological processes. Well-known examples include images of fluid flows obtained by solving the Navier–Stokes equations, skin patterning generated by Turing's celebrated diffusion–reaction model, and self-similar patterns of leaves or natural landscapes simulated by iterated function systems. This approach to image formation is most frequently applied in computer gra-phics.

Random fields. Images are modeled as samples drawn from certain random fields. The primary goal of random field modeling is to understand the probability distribution function $p(u)$, especially when Ω is a matrix of digital pixels. Classic models are inspired mainly by Gibbs fields in statistical mechanics, in which local energy constraints are imposed in the same way as in Ising crystals [7]. Random fields can also be learned from an image database through such techniques as filtering and nonpara-metric estimation by the maximum entropy principle [11].

Function spaces. In this deterministic approach, appropriate function spaces are used to calibrate image regularities, measured in some energy $E[u]$. Classic Fourier and spectral methods assume that images are drawn from $L^2(\Omega)$ and that $E[u] = \|u\|_{L^2} = \|\hat{u}\|_{L^2}$, where $\hat{u}$ is the Fourier transform. Linear filtering theory assumes that images belong to the Sobolev space $W^{1,2}(\Omega)$ and that their visual contents are measured by $E[u] = \|\nabla u\|_{L^2}$. To acknowledge the importance of edges in human visual perception [8], Rudin, Osher, and Fatemi proposed the bounded variation (BV) image model; $E[u] = \int_\Omega |Du|$ is the total variation Radon measure [10]. BV or more general Besov images have been extensively studied in wavelet theory as well. To explicitly single out edges, Mumford and Shah [9] proposed the well-known free boundary model for piecewise-smooth images:

$$E[u,\Gamma] = E[u|\Gamma] + E[\Gamma] =$$

$$\frac{\mu}{2}\int_{\Omega\backslash\Gamma} |\nabla u|^2 \, dx + \alpha H^1(\Gamma),$$

where Γ denotes the jump set and H^1 the one-dimensional Hausdorff measure, which is often replaced by length Γ in computation.

Geometric Image Models in Inpainting

We conclude with a discussion of inpainting models built on geometric image models, such as BV and Mumford–Shah images. Our first inpainting model is based on BV images [3]. The complete inpainting model is to minimize

$$E_{\text{tv}}[u|u_0,G] = \alpha\int_\Omega |Du| +$$

$$\frac{1}{2\sigma^2|\Omega\setminus G|}\int_{\Omega\backslash G}(Ku-u_0)^2 dx.$$

The noise variance σ^2 can be stat-istically estimated, and there is only one tunable constant, α. Define

$$\lambda_G(x) = \frac{1}{\sigma^2|\Omega\setminus G|}1_{\Omega\backslash G}(x),$$

the so-called inpainting mask. The model then becomes

$$E_{\text{tv}}[u|u_0,G] =$$

$$\alpha\int_\Omega |Du| + \frac{1}{2}\int_{\Omega\backslash G}\lambda_G(x)(Ku-u_0)^2\,dx,$$

which is very similar to the original Rudin–Osher–Fatemi restoration model, differing only in that the Lagrange multiplier λ is replaced by a fidelity function $\lambda_G(x)$. Figure 1shows typical output from this model.

Existence is guaranteed, but uniqueness is not [2], which should not be seen as a flaw because it models the uncertainty in Bayesian decision processes. Numerically, the model is solved by computational PDEs. The formal first variation of $E_{tv}[u]$ leads to

$$-\alpha\nabla\cdot\left[\frac{\nabla u}{|\nabla u|}\right]+K^{*}\lambda_G\left(Ku-u_0\right)=0,$$

where K^* denotes the adjoint. This degenerate nonlinear elliptic-type equation is then solved by viscosity approximation and linearization techniques.

The BV image model is geometry-motivated in that the first-order information, i.e., the length, is incorporated, as manifest in the co-area formula

$$\int_{\Omega}|Du|=\int_{-\infty}^{\infty}\mathrm{Per}\left(u<\lambda\right)d\lambda=: \int_{-\infty}^{\infty}\mathrm{length}\left(u\equiv\lambda\right)d\lambda.$$

Here the second equality is for regular functions only—a clear demonstration that total variation is a clever way to sum up the lengths of *all* level sets.

The next inpainting model, based on Mumford and Shah's object-edge model (see Figure 2), seeks to minimize the inpainting energy

$$E_{ms}[u,\Gamma|u_0,G] = E[u|\Gamma] + E[\Gamma] + E[u_0|u,G],$$

where all the terms are as discussed earlier. This free boundary model has a nice Γ-convergence approximation, in which Γ, usually a computational headache, is approximated by a signature function $z(x)$ on Ω. z is 1 almost everywhere, except along a narrow band of Γ, where it drops sharply (depending on a small control parameter ε) to 0. Esedoglu and Shen [6] have shown that inpainting provides a perfect market for Γ-convergence approximation: Unlike segmentation, inpainting seeks only the ideal image u, not Γ or z. Γ-convergence approximation substantially lessens the computational burden by reducing the original free boundary Euler–Lagrange equations to a coupled system of two well-behaved elliptic equations [6].

Figure 1. *Bounded variation inpainting. Top, a blurry image with 60 packets randomly lost during the transmission process; bottom, after deblurring and error concealment by TV inpainting.*

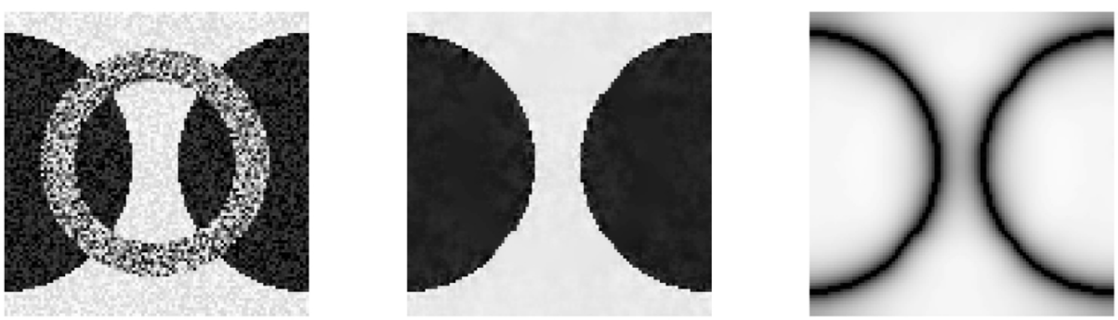

Figure 2. *Mumford–Shah inpainting for disoccluding two disks occluded by a ring. Left, noisy image to be inpainted; center, inpainting output* u; *right, inpainting output* z.

Both BV and Mumford–Shah images consider only the first-order geometry of level sets or edges, and are often sufficient for classic tasks like restoration and segmentation. For inpainting, we have demonstrated that high-order geometric information like curvature is necessary to avoid visual defects. The key tool is Euler's elastica curve model

$$e[\gamma]=\int_{\lambda}\left(a+b\kappa^2\right)dx,$$

where κ denotes the curvature of a curve γ. Originally studied by Euler in 1744 to model a one-dimensional elastic rod and later employed by Birkhoff and de Boor as a nonlinear spline model, it was first introduced into computer vision by Mumford. By formally imposing $e[\gamma]$on all the level sets, we obtain the so-called elastica image model [2]:

$$E[u] = \int_\Omega \left(a + b\kappa^2\right)|Du|,$$
$$\kappa = \nabla \cdot \left[\frac{\nabla u}{|\nabla u|}\right].$$

Similarly, by replacing the length energy in the Mumford–Shah image model by $e[\Gamma]$, we obtained the Mumford–Shah–Euler image model [2]. Except for some progress from the De Giorgi school, theoretical knowledge of these two models is still quite limited. The associated formal Euler–Lagrange PDEs do come with the desired properties, such as nonlinear transport and curvature-driven diffusion, which are essential for inpainting from the PDE point of view [4]. Numerical results based on nonlinear PDEs and high-order Γ-convergence approximation also confirm their advantages for faithful inpainting.

Acknowledgments

I would like to thank Gil Strang, Tony Chan, and Stan Osher for their invaluable teaching and inspiration; most of the ideas described here were developed during a CAM assistant professorship at UCLA, under the supervision of Tony Chan and Stan Osher. I also thank David Mumford for a visiting scientist opportunity with Brown's Pattern Theory Group in the spring of 2000, during which I gradually formed a broader view of image processing, as presented in this article.

The work described here was supported by the National Science Foundation, grant DMS–0202565.

References

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

[2] T.F. Chan, S.-H. Kang, and J. Shen, *Euler's elastica and curvature based inpaintings*, SIAM J. Appl. Math., 63–2, (2002), 564–592.

[3] T.F. Chan and J. Shen, *Mathematical models for local nontexture inpaintings*, SIAM J. Appl. Math., 62–3 (2001), 1019–1043.

[4] T.F. Chan and J. Shen, *Inpainting based on nonlinear transport and diffusion*, Amer. Math. Soc. Contemp. Math., Inverse Problems, Image Analysis, & Medical Imaging, Z. Nashed & O. Scherzer, eds., 2002.

[5] T.F. Chan, J.Shen, and L.Vese, *Variational (PDE) models in image processing*, Notices Amer. Math. Soc., 50 (2003), 14–26.

[6] S. Esedoglu and J. Shen, *Digital inpainting based on the Mumford–Shah–Euler image model*, European J. Appl. Math., 13 (2002), 353–370.

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

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

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

[10] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), 259–268.

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

Jianhong (Jackie) Shen (jhshen@math.umn) is an assistant professor of mathematics at the University of Minnesota.