Inpainting and the Fundamental Problem of Image Processing

Jianhong (Jackie) Shen *

December 24, 2002

Inpainting is an artistic synonym for image interpolation. The fundamental problem of image processing (IP) is to answer the key question "what do we mean by images." How are the two connected? Without knowing what images are, it is impossible to successfully reconstruct the missing data.

The previous article by Sapiro in SIAM News (Vol. 35, No. 4, May 2002) already gave a detailed account on the motivations and numerous applications of image inpainting. It was Sapiro's group who first borrowed the terminology "inpainting" from museum restoration artists when they developed their first 3rd order PDE based inpainting model [1]. The approach by the UCLA group has been different in that we have developed inpainting models based upon the Bayesian rationale, where the answer to the above key question in IP, or the identification of appropriate image priors, plays a critical role. Most papers mentioned here can be found in the website of UCLA's IP group at www.math.ucla.edu/~imagers.

1 Description of 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 $\mathbb{R}^2$. Due to factors such as object occlusion in visual fields and packet loss in wireless communication, there is a subset G of Ω where image data are missing or unaccessible.

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 $u_0\big|_{\Omega\setminus G}$.

Human observers experience the occlusion effect almost all the time, since most objects around are not transparent. Still the world looks perfectly ordered and integrated, rather than being a clutter of independent discrete pieces. This is our Mother Nature's answer to the inpainting problem. In vision and cognitive science, it is believed that human beings are always unawarely but cleverly applying the rules of Bayesian inferencing and decision making.

*Shen is with the School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA, Email: jhshen@math.umn.edu, Tel: (612) 625-3570. This work is supported by NSF under grant DMS-0202565. Most ideas described here were developed when the author was a CAM assistant professor at UCLA under the supervising of Tony F. Chan and Stan Osher.

2 Ingredients of Bayesian Inpainting

In the Bayesian framework, inpainting is to maximize the posterior probability: $p(u|u_0, G)$. By Bayes' formula, we have

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

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

Once u_0 and G are given, $p(u_0|G)$ is simply a normalization constant. Thus we are essentially to maximize the product of the so called *data model* $p(u_0|u, G)$ and the image *prior model* $p(u)$. A typical data model in IP is blurring followed by white noise pollution: $u_0\big|_{\Omega\setminus G} = (Ku + n)\big|_{\Omega\setminus G}$. Here K is a linear blurring operator and *lowpass*: $K1 = 1$, and n an additive Gaussian white noise.

Working with logarithm likelihood functions $E = -\frac{1}{\beta}\ln p$, we are to minimize

$$E[u|u_0, G] = E[u_0|u, G] + E[u],$$

up to an additive constant. If $\beta = 1/(kT)$ with Boltzmann constant k and absolute temperature T, then $p = \frac{1}{Z}\exp(-\beta E)$ is formally Gibbs' formula in statistical mechanics which links the energy of an ensemble to its likelihood. (Here Z is the partition function defining the free energy.)

The data model by Gaussian noise leads to

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

where $dx = dx_1 dx_2$ is the Lebesgue area element, σ^2 the variance, and $|\Omega\setminus G|$ the Lebesgue measure. Eventually, the key to the inpainting problem is to employ a suitable image prior $p(u)$ or $E[u]$, which is the fundamental problem in IP.

3 Fundamental Problem of Image Processing

What are images, mathematically speaking? Engineers have worked on IP for nearly a century without bothering to ask the question. Why should mathematicians care? In our opinion, the value of asking and answering such a fundamental question is exactly like that of laying down the Hilbert space foundation for quantum mechanics. Without a satisfactory answer, IP may never be able to become a genuine new branch of mathematics.

A general IP problem can be modeled as an input-output system [5]:

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

where the input U_0 could be a single image or an image sequence observed, $\mathcal{T}$ a linear or nonlinear image processor such as restoration and compression, and U the quested image feature. In the inpainting case, for example, $U_0 = (u_0, G)$ and

$U = u$. Knowing what class of objects U_0 and U (or the definition and range domains of $\mathcal{T}$) one is dealing with is thus crucial for effective mathematical modeling, analysis, and computation of $\mathcal{T}$.

We now describe three major approaches to the fundamental problem.

I. Physical simulation. Images are generated by simulating the underlying physical, chemical, or biological processes. Well known examples include: the images of fluid flows by solving the Navier-Stokes, Turing's celebrated diffusion-reaction model for skin pattern generation, and self-similar patterns of certain leaves or natural landscapes simulated by iterated function systems. This approach to image formation is most frequently applied in computer graphics.

II. 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. Classical models are mostly inspired by Gibbs fields in statistical mechanics, in which local energy constraints are imposed as in Ising crytals [7]. Random fields can also be learned from an image database by techniques such as filtering and non-parametric estimation by the maximum entropy principle [11].

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

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

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

4 Inpainting by Geometric Image Models

We now discuss inpainting models built upon geometric image models such as the BV and Mumford-Shah images.

Our first inpainting model is based on BV images [3]. Thus the complete inpainting model is to minimize

$$E_{\mathrm{tv}}[u|u_0, G] = \alpha \int_\Omega |Du| + \frac{1}{2\sigma^2|\Omega\backslash G|} \int_{\Omega\backslash G} (Ku - u_0)^2 dx.$$

The noise variance σ^2 can be statistically estimated and there is only one tunable constant α. Define $\lambda_G(x) = \dfrac{1}{\sigma^2|\Omega\backslash G|} 1_{\Omega\backslash G}(x)$, the so called inpainting mask. Then the model becomes

$$E_{\mathrm{tv}}[u|u_0, G] = \alpha \int_\Omega |Du| + \frac{1}{2} \int_{\Omega\backslash G} \lambda_G(x)(Ku - u_0)^2 dx,$$

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

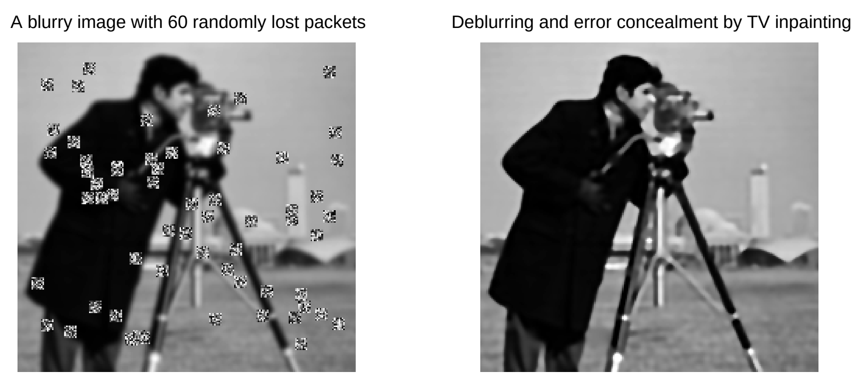

Figure 1: BV Inpainting for the error concealment of a blurry image with 60 packets randomly lost during the transmission process.

Existence is guaranteed but uniqueness is not [2], which should not be blamed because it models the uncertainty in Bayesian decision. Numerically, the model is carried out by computational PDEs. The formal first variation of $E_{\mathrm{tv}}[u]$ leads to

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

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 since the first order information, i.e. the length, is incorporated, as manifest in the co-area formula

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

Here the second equality is for regular functions only, which clearly shows that total variation is a clever way to sum up the length of *all* level sets.

The next inpainting model is based on Mumford and Shah's object-edge model, and is to minimize the inpainting energy

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

where all the terms have been discussed before. This free boundary model has a nice Γ-convergence approximation, in which the computational headache Γ 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] shows that inpainting provides a perfect market for Γ-convergence approximation, since unlike segmentation, inpainting only quests for 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].

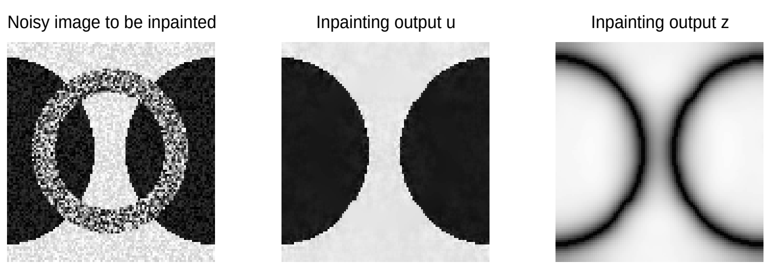

Figure 2: Mumford-Shah inpainting for disoccluding two disks occluded by a ring.

Both BV and Mumford-Shah images only consider the first order geometry of level sets or edges. They are often sufficient for classical tasks like restoration and segmentation. For inpainting, we have demonstrated that high order geometric information like curvature becomes necessary to avoid visual defects. The key tool is Euler's elastica curve model $e[\gamma] = \int_\gamma (a + b\kappa^2) dx$, where κ denotes the curvature of a curve γ. It was originally studied by Euler in 1744 to model 1-D elastic rod, later employed by Birkhoff and de Boor as a nonlinear spline model, and 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]:

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

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 some works from the De Giorgi's school, theoretical knowledge on these two models is still quite limited. Their 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. I also must thank David Mumford for a visiting scientist opportunity at Brown's Pattern Theory Group in spring 2000, during which a broader view of IP was gradually formed as presented above.

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):564–592, 2002.

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

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

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

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

[7] 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.

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

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

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

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