TOTAL VARIATION IMAGE RESTORATION WITH LOCAL CONSTRAINTS

M. Bertalmio, V. Caselles, B. Rougé and A. Solé*
Dept. of Technology, University of Pompeu-Fabra,
Passeig de Circumvalació, 8,
08003 Barcelona, Spain
Tel: +34 93 542 24 21; fax: +34 93 542 25 17
and *CNES, Av. Ed Belin, 31401 Toulouse Cedex 4, France
e-mail: vicent.caselles@tecn.upf.es, bernard.rouge@cnes.fr

ABSTRACT

L. Rudin, S. Osher and E. Fatemi proposed a method for image restoration based on the minimization of the Total Variation submitted to the constraints given by the image acquisition model. To ensure the satisfaction of the constraints, a Lagrange multiplier λ has to be estimated. We observe that if λ is global, the constraints are also globally satisfied, but not locally. We discuss the problem of imposing local constraints with a finite number of Lagrange multipliers associated to a partition of the image and we display some experimental results obtained with this model.

1 INTRODUCTION

We suppose that our image (or data) z is a scalar function defined on a rectangle Ω in $\mathbb{R}^2$. Generally, the degradation of the original image u occurs during image acquisition and can be modeled by a linear and translation invariant blur and additive noise. The equation relating u to z can be written as

$$z = Ku + n \tag{1}$$

where K is a convolution operator with impulse response k, i.e., $Ku = k * u$, and n is an additive white noise of standard deviation σ. In practice, the noise can be considered as Gaussian.

The problem of recovering u from z is ill-posed. First, the blurring operator need not be invertible. Second, if the inverse operator K^{-1} exists, applying it to both sides of (1) we obtain

$$K^{-1}z = u + K^{-1}n. \tag{2}$$

Writing $K^{-1}n$ in he Fourier domain, we have $K^{-1}n = \left(\frac{\hat{n}}{\hat{k}}\right)^{\vee}$ where $\hat{f}$ denotes the Fourier transform of f and $f^{\vee}$ denotes the inverse Fourier transform. From this equation, we see that the noise might blow up at the frequencies for which $\hat{k}$ vanishes or it becomes small.

Several methods have been proposed to recover u. Most of them can be classified as regularization methods which may take into account statistical properties (Wiener filters), information theoretic properties ([2]), a priori geometric models ([9]) or the functional analytic behavior of the image given in terms of its wavelet coefficients ([3]).

The tipical strategy to solve this ill-conditioning is regularization. Then the solution of (1) is estimated by minimizing a functional

$$J_\gamma(u) = \| Ku - z \|_2^2 + \gamma \| Qu \|_2^2 \tag{3}$$

with $\gamma > 0$, which yields the estimate $u_\gamma = (K^tK + \gamma Q^tQ)^{-1}K^tz$ where Q is a linear regularization operator. Observe that to obtain u_γ we have to solve a linear equation. The role of Q is on one hand to move the small eigenvalues of K^tK away from zero while leaving the large eigenvalues unchanged, and, on the other hand, to incorporate the a priori (smoothness) knowledge that we have on u. One of the first linear operators used was $Qu = \nabla u$. In this case the smoothness constraint is too restrictive. Indeed, functions $u \in L^2(\Omega)$ ($L^2(\Omega)$ denotes the space of square integrable functions in Ω) such that $Du \in L^2(\Omega)$ cannot have discontinuities along rectifiable curves. These observations motivated the introduction of Total Variation in image restoration models by L. Rudin, S. Osher and E. Fatemi in their seminal work [9]. The a priori hypothesis is that functions of bounded variation (the BV model) ([5]) are a reasonable functional model for many problems in image processing, in particular, for restoration problems ([9]). Tipically, functions of bounded variation have discontinuities along rectifiable curves, being continuous in some sense (in the measure theoretic sense) away from discontinuities. The discontinuities could be identified with edges. The ability of this functional to describe textures is less clear: some textures can be recovered, but up to a certain scale of oscillation.

On the basis of the BV model, Rudin-Osher-Fatemi [9] proposed to solve the following constrained minimization problem

$$\begin{aligned} &\text{Minimize} \int_\Omega |Du|\,dx \quad \text{with constraints} \\ &\int_\Omega Ku = \int_\Omega z,\ \int_\Omega |Ku - z|^2\,dx = \sigma^2|\Omega|. \end{aligned} \tag{4}$$

Here we assume that $u \in L^2(\Omega)$ is a function of bounded variation in Ω, i.e., the distributional gradient of u is a measure with finite total mass $\int_\Omega |Du|$. The first constraint corresponds to the assumption that the noise has zero mean, and the second that its standard deviation is σ. The constraints are a way to incorporate the image acquisition model given in terms on equation (1).

1.1 The Lagrange multiplier method

In practice, the above problem is solved via the following unconstrained minimization problem

$$\text{Minimize} \int_\Omega |Du|\,dx + \frac{\lambda}{2}\int_\Omega |Ku - z|^2\,dx \qquad (5)$$

for some Lagrange multiplier λ. We note that the first constraint is authomatically satisfied if $K1 = 1$ [1]. The constraint has been introduced as a penalization term. The regularization parameter λ controls the trade-off between the goodness of fit of the constraint and the smoothnes term given by the Total Variation. In this formulation, a methodology is required for a correct choice of λ [9],[1]. The most succesful analysis of the connections between (4) and (5) were given by A. Chambolle and P.L. Lions in [1]. Indeed, they proved that, under some conditions on the kernel K and the data z, both problems are equivalent for some positive value of the Lagrange multiplier λ.

The Euler-Lagrange equation of (5) is

$$\begin{cases} -\text{div}\left(\dfrac{Du}{|Du|}\right) + \lambda K^t(Ku - z) = 0 & \text{in } \Omega \\ \\ z \cdot \nu = 0 & \text{on } \partial\Omega \end{cases} \qquad (6)$$

2 LOCAL CONSTRAINTS

If, for simplicity, we consider a discrete image model, then we may write

$$z(i,j) = Ku(i,j) + n(i,j)$$

where $n(i,j) \in N(0,\sigma^2)$. Then, if I is a subset of the indexes (i,j), then

$$S_I = \frac{1}{\sigma^2}\sum_{(i,j)\in I}(z(i,j) - Ku(i,j))^2 = \frac{1}{\sigma^2}\sum_{(i,j)\in I} n(i,j)^2$$

has a χ^2 distribution with $|I|$ (the number of points in I) degrees of freedom. In particular,

$$E(S_I) = |I| \quad \text{and} \quad var(S_I) = 2|I|.$$

By the Central Limit Theorem, we have

$$P(|S_I - |I|| > x\sqrt{2|I|}) \to \frac{1}{\sqrt{2\pi}}\int_{|t|>x} e^{-t^2/2}dt, \qquad (7)$$

as $|I| \to \infty$, or

$$P(|\tfrac{1}{|I|}\textstyle\sum_{(i,j)\in I} n(i,j)^2 - \sigma^2| > x\sigma^2\sqrt{\tfrac{2}{|I|}}) \to \tfrac{1}{\sqrt{2\pi}}\textstyle\int_{|t|>x} e^{-t^2/2}dt. \qquad (8)$$

Thus, the expression $\frac{1}{|I|}\sum_{(i,j)\in I} n(i,j)^2$ represents σ^2 with an error proportional to $\frac{1}{\sqrt{|I|}}$. To obtain this estimate we need to work at the discrete level.

We conclude that, for a large region O, we should have that the local constraint

$$\int_O (Ku - z)^2 = \sigma^2|O| \qquad (9)$$

should hold approximately. We shall check experimentally that this is not the case for the solution u^λ of (5) with λ chosen so that $\int_\Omega (Ku^\lambda - z)^2 = \sigma^2|\Omega|$. Thus, the constraints have to be imposed locally.

2.1 Local Selection of the Lagrange multiplier

Suppose that $\{O_1, ..., O_p\}$ is a partition of Ω and assume that O_i have Lipschitz boundaries. Given $\lambda_1, ..., \lambda_p$, we consider the following problem

$$\text{Minimize} \int_\Omega |Du|\,dx + \frac{1}{2}\int_\Omega \lambda(x)|Ku - z|^2\,dx \qquad (10)$$

where $\lambda(x) = \sum_{i=1}^p \lambda_i \chi_{O_i}$. One can prove that the minimization problem (10) has a solution for any set of values $\lambda_i \geq 0$ and there are values $(\lambda_1^*, ..., \lambda_p^*)$, $\lambda_i^* \geq 0$, such that there is solution of (10) which is also a solution of

$$\begin{array}{l} \text{Minimize} \displaystyle\int_\Omega |Du|\,dx \\ \text{with} \quad \displaystyle\int_{O_i} |Ku - z|^2\,dx \leq \sigma^2|O_i|,\ \forall i = 1, ..., p. \end{array} \qquad (11)$$

Moreover, for each i such that $\lambda_i^* > 0$ the constraint is satisfied with equality instead of inequality. If $\lambda_i > 0$, $i = 1, ..., p$, and K is injective then the solution is unique.

3 NUMERICAL EXPERIMENTS

The convolution kernel of the first model is given by

$$\begin{array}{l} \hat{k}(\xi,\eta) = \\ e^{-2\gamma_\xi|\xi| - 2\gamma_\eta|\eta|}\left(\frac{sin(2\pi\xi)}{2\pi\xi}\right)\left(\frac{sin(2\pi\eta)}{2\pi\eta}\right)\left(\frac{sin(\pi\xi)}{\pi\xi}\right), \end{array} \qquad (12)$$

where $\xi, \eta \in [-\frac{1}{2}, \frac{1}{2}]$, $\gamma_\xi = 1.505$, $\gamma_\eta = 1.412$. Observe that K is injective.

We have assumed that the noise is a white Gaussian noise of standard deviation σ. Following [8], the signal to noise ratio, or SNR, can be approximated by the quotient between the average of the signal and the standard deviation of the signal for the average luminance $SNR = \frac{Average}{\sigma}$. Thus $\sigma = \frac{Average}{SNR}$. In practice, the values of $SNR = 50, 100$ are realistic and give rise to the same difficulty as the real noise.

To solve (5) with a fixed value of λ we use a gradient descent method. The correct value of λ can be computed by bisection. Another possibility is to observe that the quantity $Q_\Omega(u^\lambda) = \frac{1}{|\Omega|}\int_\Omega (Ku^\lambda - z)^2$ is decreasing in

Figure 1: Original image.

Figure 2: Degraded image.

λ ([1]) and we can increase/decrease λ until the constraint $Q_\Omega(u^\lambda) = \sigma^2$ is satisfied. Figure 1 displays the original uncorrupted image. Figure 2 has been obtained by convolving Figure 1 with kernel k and adding a noise of standard deviation $\sigma = 1$. Figure 3, Top, is the restored image obtained by applying (5) with the value of λ selected so that $Q_\Omega(u^\lambda) = \sigma^2$, indeed, with $\lambda = 0.35$. In practice, one would select a value of λ which gives a visually satisfactory result, even if the constraint is not satisfied. The purpose of the experiment is to show that this value of λ does not imply that the constraint holds locally for the regions of a partition of the image. We have computed a partition using Mumford-Shah's segmentation algorithm and we display in Figure 3, Bottom, the values of $Q_O(u^\lambda) = \frac{1}{|O|}\int_O (Ku^\lambda - z)^2$ when O is the region depicted in white in Figure 4. We see that the constraint is not satisfied for this region, indeed $Q_O(u^\lambda)$ does not decrease to $\sigma^2 = 1$.

Let us describe the algorithm to minimize 10. We start with a partition of the image obtained with Mumford-Shah's segmentation algorithm.

(i) For each set of values $\lambda_i > 0$, $i = 1, ..., P$, we solve iteratively (10), until we reach the asymptotic state $u^{(\lambda_1,...,\lambda_p)}$.

(ii) Initially, we take the values of $\lambda_i > 0$ small enough so that (*) $\frac{1}{|O_i|}\int_{O_i}(Ku^\lambda - z)^2 > \sigma^2, \quad \forall i = 1, ..., p$ holds.

(iii) For each $i \in \{1, ..., p\}$ for which (*) holds, increase the value of λ_i and compute again the asymptotic state.

Figure 5 displays the restored image u obtained by using the above algorithm. In Figures 6, 7 we display the values of $Q_O(u)$ for the regions depicted in Figure 4. We may conclude that imposing local constraints enables us to get a better reconstruction both at textured and non-textured areas.

Acknowledgements. This work was supported by the

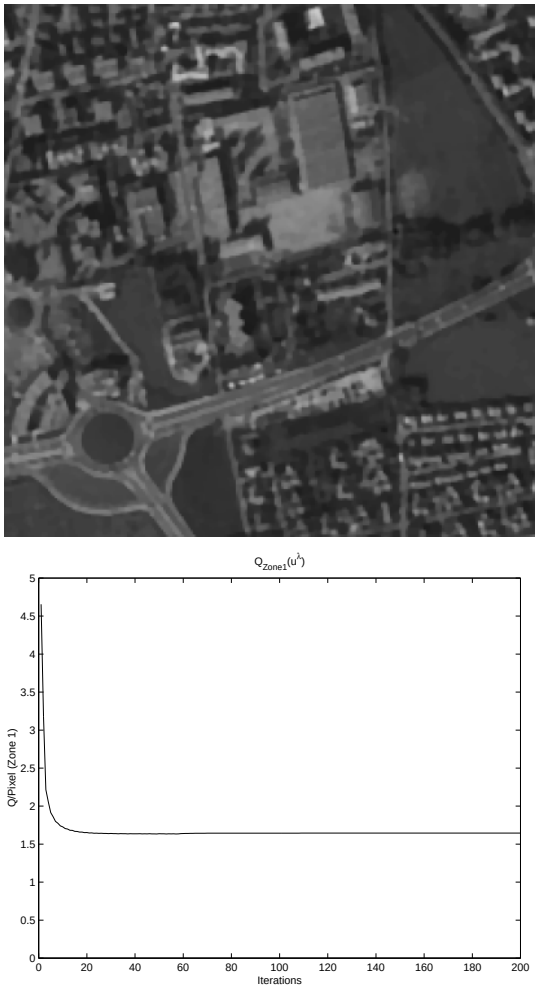

Figure 3: Top: Restored with global constraint. Bottom: Evolution of $Q_\Omega(u)$.

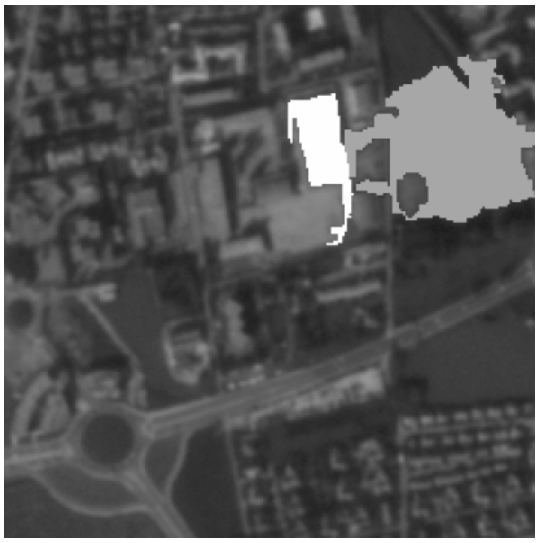

Figure 4: Two regions obtained with Mumford-Shah's algorithm: a textured region in white and a non-textured one, in light gray.

Figure 5: Restoration result with local constraints.

CNES.

References

[1] A. Chambolle and P. L. Lions, *Image Recovery via Total Variation Minimization and Related Problems*, Num. Mat. 76 (1997), no. 2, 167-188.

[2] G. Demoment, *Image reconstruction and restoration: Overview of common estimation structures and problems,* IEEE Trans. on Acoust., Speech and Signal Proc., Vol. 37, 12 (1989), pp. 2024-2036.

[3] D.L. Donoho, I.M. Johnstone, G. Kerkyacharian and D. Picard, *Wavelet Shrinkage : Asymptopia ?.* J.R. Statist. Soc. B **57** (1995), 301–369.

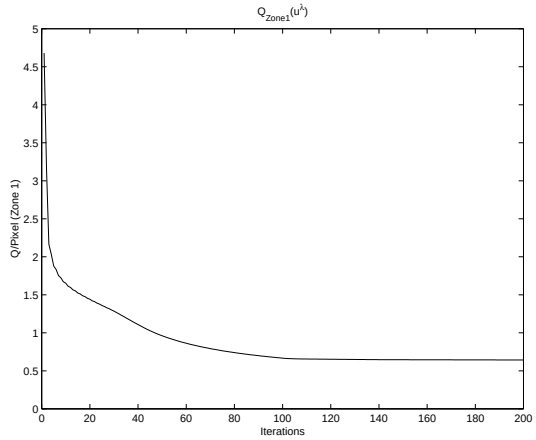

Figure 6: Evolution of $Q_O(u)$ for the white region.

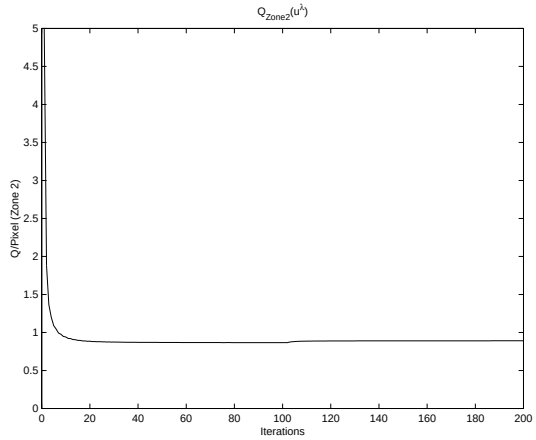

Figure 7: Evolution of $Q_O(u)$ for the light gray region.

[4] S. Durand, F. Malgouyres and B. Rougé, *Image Deblurring, Spectrum Interpolation and Application to Satellite Imaging,* Mathematical Modelling and Numerical Analysis, 1999.

[5] L.C. Evans and R.F. Gariepy, *Measure Theory and Fine Properties of Functions*, Studies in Advanced Math., CRC Press, 1992.

[6] D. Geman and G. Reynolds, *Constrained Image Restoration and Recovery of Discontinuities,* IEEE Trans. Pattern Anal. Machine Intell., 14 (1992), pp. 367-383.

[7] L. Moisan, *Extrapolation de spectre et variation totale ponderée,* Preprint 2001.

[8] B. Rougé, *Théorie de l'echantillonage et satellites d'observation de la terre.* Analyse de Fourier et traitement d'images, Journées X-UPS 1998.

[9] L.I. Rudin, S. Osher and E. Fatemi, *Nonlinear Total Variation Based Noise Removal Algorithms.* Physica D, **60**, 259–269.