International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:5, No:6, 2011

Globally convergent edge-preserving reconstruction with contour-line smoothing

Marc C. Robini, Pierre-Jean Viverge, Yuemin Zhu and Jianhua Luo

Abstract—The standard approach to image reconstruction is to stabilize the problem by including an edge-preserving roughness penalty in addition to faithfulness to the data. However, this methodology produces noisy object boundaries and creates a staircase effect. The existing attempts to favor the formation of smooth contour lines take the edge field explicitly into account; they either are computationally expensive or produce disappointing results. In this paper, we propose to incorporate the smoothness of the edge field in an implicit way by means of an additional penalty term defined in the wavelet domain. We also derive an efficient half-quadratic algorithm to solve the resulting optimization problem, including the case when the data fidelity term is non-quadratic and the cost function is non-convex. Numerical experiments show that our technique preserves edge sharpness while smoothing contour lines; it produces visually pleasing reconstructions which are quantitatively better than those obtained without wavelet-domain constraints.

I. Introduction

We consider the classical inverse problem of recovering a piecewise-smooth image $x^\star \in \mathbb{R}^K$ from some data $d \in \mathbb{R}^{K'}$ of the form

$$d = \mathcal{H}(x^\star) + e_\eta, \tag{1}$$

where $\mathcal{H} \in \mathcal{L}_{K,K'}(\mathbb{R})$ models the acquisition process—$\mathcal{L}_{m,n}(\mathbb{R})$ denotes the set of linear maps from $\mathbb{R}^m$ to $\mathbb{R}^n$—and where the noise term $e_\eta \in \mathbb{R}^{K'}$ is a realization of a zero-mean random vector η. Throughout this paper, we make no distinction between a real-valued image with K pixels and its vector representation in $\mathbb{R}^K$.

A common estimate of the original image $x^\star$ is defined as any global minimum of a cost function $U : \mathbb{R}^K \to \mathbb{R}$ which combines a data-fidelity term Φ_d with a stabilization term Ψ weighted by a parameter $\lambda > 0$:

$$U(x) = \Phi_d(x) + \lambda\Psi(x). \tag{2}$$

The justification for this choice lies within the regularization framework [1] or its Bayesian interpretation [2].

The data-fidelity term accounts for the noise characteristics and is usually defined by

$$\Phi_d(x) = \sum_{k\in[\![1,K']\!]} \phi_k\big(|\mathcal{H}_k(x) - d_k|\big), \tag{3}$$

where the ϕ_k's are increasing in $\mathbb{R}_+$ ($[\![1,K']\!]$ is a shorthand notation for the set $\{1,\dots,K'\}$). In the Bayesian setting, this form is obtained under the assumption that the noise components $\eta_1,\dots,\eta_{K'}$ are independent and that, for each k, η_k has a density proportional to $\exp(-\phi_k(|t|))$.

The regularization functional Ψ is intended to promote the formation of smooth regions separated by edges; it is of the form

$$\Psi(x) = \sum_{l\in[\![1,L]\!]} \psi_l\big(\|\mathcal{R}_l(x)\|\big), \tag{4}$$

where $\|\cdot\|$ is the ℓ_2-norm, $\mathcal{R}_l \in \mathcal{L}_{K,\rho_l}(\mathbb{R})$ ($\rho_l \in \mathbb{N}^*$), and the functions ψ_l are increasing in $\mathbb{R}_+$. Most of the time, the functions ψ_l are the same for all l and $\{\mathcal{R}_l : l \in [\![1,L]\!]\}$ is a set of first-order difference operators ($\rho_l = 1$) or a discrete approximation to the gradient ($\rho_l = 2$). The ψ_l's are called *potential functions* (PFs) in the Bayesian framework. On the one hand, some authors encourage the use of convex PFs which ensure the convexity of U and reduce smoothing in the vicinity of discontinuities under mild conditions [3]. On the other hand, one can be interested in non-convex PFs which yield sharper edges [4] at the expense of increased optimization difficulty.

The formulation of the image reconstruction problem as the minimization of a cost function of the form (2)–(4) has proven effective, and yet it has two limitations.

1) Noisy contour lines: Regularization functionals of the type of (4) do not embed prior knowledge on the geometry of edges and thus tend to produce noisy object boundaries which are not faithful to the original image. The accepted way to reduce this effect is to take into account the mutual dependence between neighboring discontinuities by introducing an explicit edge process, either boolean [5], [6] or continuous [7], [8]. However, boolean edge processes drastically increase computational complexity and continuous edge processes give disappointing results.

2) Staircase effect: Convex edge-preserving PFs are often similar to the identity on $\mathbb{R}_+$, but with zero derivative at the origin to make the regularization term Ψ differentiable. A well-known example is the approximation $\mathrm{id}^\delta_{\mathbb{R}_+}$ to $\mathrm{id}_{\mathbb{R}_+}$ defined by

$$\mathrm{id}^\delta_{\mathbb{R}_+}(t) = \sqrt{\delta^2 + t^2} - \delta, \tag{5}$$

where $\delta > 0$ is small compared to the expected amplitude range of the original image. When $\psi_l = \mathrm{id}_{\mathbb{R}_+}$, $\Psi(x)$ is the discrete total variation (TV) of x. Despite its popularity, TV regularization produces blocky images (see, e.g., [9], [10]) and so do PFs close to the identity. This phenomenon is called the staircase effect because ramps (affine regions) tend to turn into

M. Robini and Y. Zhu are with CREATIS (CNRS research unit UMR5220 and INSERM research unit U1044), INSA-Lyon, 69621 Villeurbanne, France (e-mail: marc.robinicreatis.insa-lyon.fr).

P.-J. Viverge is with the electrical engineering department of the INSA-Lyon, 69621 Villeurbanne, France.

J. Luo is with the College of Life Science and Technology, Shanghai Jiao Tong University, Shanghai 200240, China.

This work was partly supported by the French ANR (ANR-09-BLAN-0372-01) and by the Chinese NSFC (30911130364).

stairs (piecewise-constant regions). The staircase effect is also observed for non-convex PFs and is more pronounced in this case (see, e.g., [11]).

In this paper, we propose an improved model constructed by adding constraints in a multiresolution space. More precisely, we consider the minimization of an augmented functional $V : \mathbb{R}^K \to \mathbb{R}$ defined by

$$V(x) = U(x) + \widetilde{\lambda}\,(\widetilde{\Psi} \circ \mathcal{T})(x), \tag{6}$$

where $\mathcal{T}$ is a multiresolution directional transform (e.g., wavelet, curvelet), $\widetilde{\Psi}$ is a smoothness penalty term operating on the detail coefficients, and $\widetilde{\lambda} > 0$ weights the influence of $\widetilde{\Psi}$. Unlike other edge-continuation methods, this technique incorporates the smoothness features of the edge field implicitly rather than explicitly. Not only does it encourage the formation of smooth contours while preserving edges, but it also reduces the staircase effect. We also derive an efficient fixed-point iteration scheme to minimize V; the associated convergence results extend those presented in [12].

The paper is organized as follows. In Section II, we discuss our approach to the issue of edge continuation. Section III is devoted to the description of the optimization algorithm along with its convergence properties. Experimental results are given in Section IV, followed by concluding remarks.

II. Implicit Edge Continuation

A. General description

We start with the assumption that the level curves of the underlying continuous original image (from which $x^\star$ is sampled) are smooth except at finitely many points. In a multiscale directional transform domain, this prior information is synonymous with clean and strongly oriented detail images. Consequently, edge continuation can be incorporated into the reconstruction process by means of an additional penalty term of the form

$$\widetilde{\Psi}(\widetilde{x}) = \sum_{m \in [\![1,M]\!]} \widetilde{\psi}_m\big(|\mathcal{D}_m(\widetilde{x})|\big), \tag{7}$$

where $\widetilde{x}$ denotes the multiresolution decomposition of x, the $\mathcal{D}_m$'s are discrete directional derivatives operators acting on the detail coefficients, and the functions $\widetilde{\psi}_m$ are increasing in $\mathbb{R}_+$.

Let $\mathcal{T} \in \mathcal{L}_{K,\widetilde{K}}(\mathbb{R})$ be the considered multiresolution transform. Our estimate of $x^\star$ is defined as any global minimum of the augmented cost function $V : \mathbb{R}^K \to \mathbb{R}$ obtained by combining the contour-line smoothing term $\widetilde{\Psi}$ with the standard cost function U:

$$\begin{aligned} V(x) &= \Phi_d(x) + \lambda\Psi(x) + \widetilde{\lambda}\,(\widetilde{\Psi} \circ \mathcal{T})(x) \\ &= \sum_{k \in [\![1,K']\!]} \phi_k\big(|\mathcal{H}_k(x) - d_k|\big) \\ &\quad + \lambda \sum_{l \in [\![1,L]\!]} \psi_l\big(\|\mathcal{R}_l(x)\|\big) \\ &\quad + \widetilde{\lambda} \sum_{m \in [\![1,M]\!]} \widetilde{\psi}_m\big(|(\mathcal{D}_m \circ \mathcal{T})(x)|\big), \end{aligned} \tag{8}$$

where the positive parameters λ and $\widetilde{\lambda}$ control the degree of smoothing in the space domain and in the transformed domain, respectively. The interest of this approach are that (i) it avoids the delicate use of explicit-line processes such as proposed in [5]–[8], (ii) it involves only one extra hyper-parameter (namely, $\widetilde{\lambda}$), and (iii) the associated increase in optimization complexity is small if the $\widetilde{\psi}_m$'s are convex.

In the following subsection, we refine the description of the contour-line smoothing term $\widetilde{\Psi}$ in the case of a non-redundant wavelet transform. We focus on this situation for computational efficiency reasons and because the construction is easier when directional selectivity is limited. Yet, the approach can be applied to the case of a redundant transform with high directional selectivity (e.g., the curvelet transform) at the expense of additional design efforts and increased computational burden.

B. Wavelet-domain edge continuation

Let $J \in \mathbb{N}^*$. The J-level decomposition of an image $x \in \mathbb{R}^K$ in an orthonogonal (or biorthogonal) wavelet basis is a set of subimages

$$\big\{\, \mathcal{A}_0(x)\,,\, \mathcal{A}_j^{\mathrm{h}}(x)\,,\, \mathcal{A}_j^{\mathrm{v}}(x)\,,\, \mathcal{A}_j^{\mathrm{d}}(x) : j \in [\![0, J-1]\!] \,\big\},$$

where the index j represents the scale, $\mathcal{A}_0(x)$ is the approximation of x at resolution 2^{-J} (the resolution of x is equal to one and its scale is zero), and the detail images $\mathcal{A}_j^{\mathrm{h}}(x)$, $\mathcal{A}_j^{\mathrm{v}}(x)$ and $\mathcal{A}_j^{\mathrm{d}}(x)$ convey the difference of information between $\mathcal{A}_j(x)$, the approximation of x at resolution 2^{j-J}, and the finer approximation $\mathcal{A}_{j+1}(x)$ ($\mathcal{A}_j^{\mathrm{h}}$ gives the horizontal high frequencies, that is, the vertical edges, $\mathcal{A}_j^{\mathrm{v}}$ gives the vertical high frequencies, that is, the horizontal edges, and $\mathcal{A}_j^{\mathrm{d}}$ gives the high frequencies in both directions). The wavelet transform $\mathcal{T}$ can be viewed as the concatenation of the component functions of the maps $\mathcal{A}_0 \in \mathcal{L}_{K,4^{-J}K}(\mathbb{R})$ and $\mathcal{A}_j^{\mathrm{h}}, \mathcal{A}_j^{\mathrm{v}}, \mathcal{A}_j^{\mathrm{d}} \in \mathcal{L}_{K,4^{j-J}K}(\mathbb{R})$ for $j \in [\![0, J-1]\!]$.

Smooth level curves in the original continuous image translate to the fact that the horizontal and vertical detail images of $x^\star$ are respectively vertically- and horizontally-oriented. This suggests to define the contour-line smoothing term $\widetilde{\Psi}_{\mathrm{w}}$ as follows:

$$\begin{aligned} (\widetilde{\Psi}_{\mathrm{w}} \circ \mathcal{T})(x) = \sum_{j \in [\![0,J-1]\!]} \Bigg(& \sum_{k \in [\![1,K_j^{\mathrm{v}}]\!]} \widetilde{\psi}\big(\mu_j \big|(\mathcal{D}_{j,k}^{\mathrm{v}} \circ \mathcal{A}_j^{\mathrm{h}})(x)\big|\big) \\ & + \sum_{k \in [\![1,K_j^{\mathrm{h}}]\!]} \widetilde{\psi}\big(\mu_j \big|(\mathcal{D}_{j,k}^{\mathrm{h}} \circ \mathcal{A}_j^{\mathrm{v}})(x)\big|\big) \Bigg) \end{aligned} \tag{9}$$

where $\widetilde{\psi}$ is increasing in $\mathbb{R}_+$, the μ_j's are positive reals, and the $\mathcal{D}_j^{\mathrm{v}}$'s and the $\mathcal{D}_j^{\mathrm{h}}$'s are vertical and horizontal discrete derivative operators, respectively. The outer summation is over the scale and the inner summations are over the locations in the detail images at which the derivatives are computed. A natural choice for $\widetilde{\psi}$ is the identity on $\mathbb{R}_+$, as we would ideally like to preserve both sharp transitions and smooth variations along the preferential directions of the detail images. However, we advocate using a differentiable convex function similar to $\mathrm{id}_{\mathbb{R}_+}$ but with zero derivative at the origin (an example of which

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:5, No:6, 2011

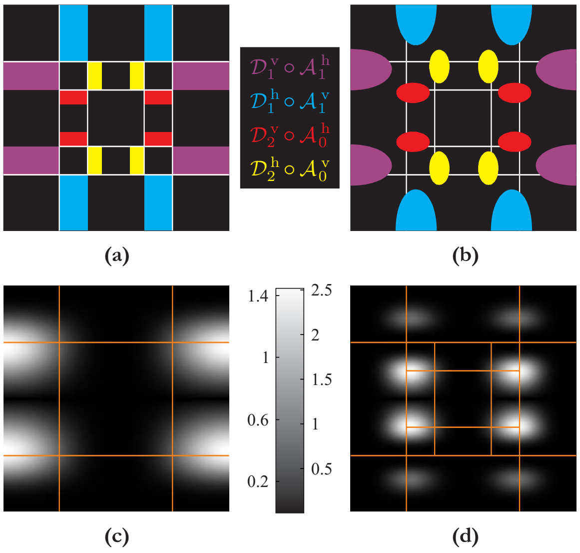


Fig. 1. Frequency regions penalized by the wavelet-domain smoothing term $\widetilde{\Psi}_{\mathrm{w}}$ for $J = 2$: (a) schematization of the passbands of the involved filtering operators; (b)–(d) actual 3 dB passbands and frequency responses of $\mathcal{D}_1^{\mathrm{v}} \circ \mathcal{A}_1^{\mathrm{h}}$ and $\mathcal{D}_0^{\mathrm{v}} \circ \mathcal{A}_0^{\mathrm{h}}$ for biorthogonal spline wavelets and first-order differences.

is given by (5)) so as to facilitate the optimization process without compromising the quality of the reconstructions. The weights μ_j allow to strengthen the edge continuation effect at some particular resolution level(s).

From a filtering perspective, the action of $\widetilde{\Psi}_{\mathrm{w}}$ is to discard the solutions that convey significant information in the frequency regions schematized in Fig. 1(a). These regions are exemplified in Figs. 1(b)–(d) for biorthogonal spline wavelets with two vanishing moments and when $\mathcal{D}_j^{\mathrm{v}}$ and $\mathcal{D}_j^{\mathrm{h}}$ compute differences between neighboring detail coefficients, as used in our experiments. It appears that $\widetilde{\Psi}_{\mathrm{w}}$ not only favors the formation of horizontal and vertical edges, but also preserves diagonal boundaries, which is why we do not add a penalty on the diagonal subbands $\mathcal{A}_j^{\mathrm{d}}(x)$, $j \in [\![0, J-1]\!]$.

III. Deterministic Relaxation

In this section, we focus on the minimization of the augmented cost function V given in (8) (the proposed wavelet-domain smoothing term $\widetilde{\Psi}_{\mathrm{w}}$ defined by (9) is an instance of (7)). To simplify notation, we make no distinction between a linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$ and its $n \times m$ matrix representation and between a vector of $\mathbb{R}^n$ and its $n \times 1$ column matrix representation. For further convenience, we write V in the more compact form

$$V(x) = \sum_{k \in [\![1,K']\!]} \phi_k\big(|(\mathcal{H}x - d)_k|\big) + \sum_{l \in [\![1,L']\!]} \theta_l\big(\|\mathcal{Q}_l x\|\big) \quad (10)$$

where $L' = L + M$ and

$$(\theta_l, \mathcal{Q}_l) = \begin{cases} (\lambda\,\psi_l, \mathcal{R}_l) & \text{if } l \in [\![1, L]\!], \\ (\widetilde{\lambda}\,\widetilde{\psi}_{l-L}, \mathcal{D}_{l-L}\mathcal{T}) & \text{if } l \in [\![L+1, L']\!]. \end{cases} \quad (11)$$

A. Construction of the algorithm

We assume that each function $\theta_l : \mathbb{R}_+ \to \mathbb{R}$ satisfies the following groups of conditions.

C1. θ_l is increasing and C^1, and $\theta_l(0) = 0$.

C2. θ_l is four times differentiable at zero and $\theta_l'''(0) = 0$.

C3. The function

$$\theta_l^\dagger : t \in \mathbb{R}_+ \mapsto \begin{cases} \theta_l'(t)/t & \text{if } t > 0, \\ \theta_l''(0) & \text{if } t = 0, \end{cases} \quad (12)$$

is continuous and strictly decreasing, and $\lim_{t\to+\infty} \theta_l^\dagger(t) = 0$.

(Conditions C1 are standard in regularized reconstruction, conditions C2 are technical requirements for convergence proofs, and conditions C3 guarantee edge-preservation properties.) For the functions ϕ_k, we define $\phi_k^\dagger$ similarly to $\theta_l^\dagger$, that is, $\phi_k^\dagger(t) = \phi_k'(t)/t$ if $t > 0$ and $\phi_k^\dagger(0) = \phi_k''(0)$, and we assume that either C4 or C4$'$ below holds.

C4. The ϕ_k's satisfy the same conditions as the θ_l's.

C4$'$. $\phi_k(t) = t^2$ for all k.

For any $x \in \mathbb{R}^K$, we define the diagonal matrices

$$\mathcal{E}_\phi(x) = \operatorname{diag}\big(\phi_k^\dagger\big(|(\mathcal{H}x - d)_k|\big) : k \in [\![1, K']\!]\big) \quad (13)$$

$$\text{and} \quad \mathcal{E}_\theta(x) = \operatorname{diag}\big(\theta_l^\dagger\big(\|\mathcal{Q}_l x\|\big)\,\mathbb{1}_{\rho_l} : l \in [\![1, L']\!]\big), \quad (14)$$

where $\mathbb{1}_r = (1, \dots, 1) \in \mathbb{R}^r$ and ρ_l is the number of rows of $\mathcal{Q}_l$. The first-order necessary condition for $x \in \mathbb{R}^K$ to be a local minimum of V (namely $\nabla V(x) = 0$) writes

$$\underbrace{\big(\mathcal{H}^T\mathcal{E}_\phi(x)\mathcal{H} + \mathcal{Q}^T\mathcal{E}_\theta(x)\mathcal{Q}\big)}_{=:\,\mathcal{M}(x)} x = \mathcal{H}^T\mathcal{E}_\phi(x)d, \quad (15)$$

where $\mathcal{Q} = [\,\mathcal{Q}_1^T, \dots, \mathcal{Q}_{L'}^T\,]^T$ (i.e., $\mathcal{Q}$ is the vertical concatenation of the matrices $\mathcal{Q}_l$). For any $x \in \mathbb{R}^K$, the symmetric matrix $\mathcal{M}(x)$ is positive semidefinite; it is positive definite if and only if

C5. $\ker(\mathcal{H}) \cap \ker(\mathcal{Q}) = \{0\}$.

If C5 is satisfied, (15) suggests the following iterative relaxation algorithm starting from a given $x^{(0)} \in \mathbb{R}^K$:

$$x^{(n+1)} = \big(\mathcal{M}(x^{(n)})\big)^{-1}\mathcal{H}^T\mathcal{E}_\phi(x^{(n)}). \quad (16)$$

This fixed-point iteration can be equivalently written as

$$x^{(n+1)} = \arg\min_{y \in \mathbb{R}^K} V_0\big(y, e_\phi(x^{(n)}), e_\theta(x^{(n)})\big), \quad (17)$$

where $V_0 : \mathbb{R}^K \times \mathbb{R}^{K'} \times \mathbb{R}^{L'} \to \mathbb{R}$ is defined by

$$V_0(x, \delta, \varepsilon) = \sum_{k \in [\![1,K']\!]} \delta_k(\mathcal{H}x - d)_k^2 + \sum_{l \in [\![1,L']\!]} \varepsilon_l\|\mathcal{Q}_l x\|^2, \quad (18)$$

and where $e_\phi = (e_{\phi,1}, \dots, e_{\phi,K'})$ and $e_\theta = (e_{\theta,1}, \dots, e_{\theta,L'})$ are the vector-valued functions with components

$$e_{\phi,k}(x) = \phi_k^\dagger\big(|(\mathcal{H}x - d)_k|\big) \ \text{ and } \ e_{\theta,l}(x) = \theta_l^\dagger\big(\|\mathcal{Q}_l x\|\big). \quad (19)$$

The map $x \mapsto V_0(x, \delta, \varepsilon)$ is a positive definite quadratic form for all (δ, ε), and thus (16) belongs to the class of half-quadratic algorithms.

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:5, No:6, 2011

B. *Convergence properties*

The convergence properties of the proposed optimization algorithm (16) are given in Theorems 1 and 2 below. The proof of Theorem 1 uses ideas from [12] (it is omitted because of space limitations), and Theorem 2 is a corollary of Theorem 1. We denote by Ξ the set of critical points of V, that is,

$$\Xi \;=\; \left\{x \in \mathbb{R}^K \,\middle|\, \nabla V(x) = 0\right\} \tag{20}$$

and we use the following terminology: a sequence $(x^{(n)})_n$ in $\mathbb{R}^K$ is said to be *generated by* (16) if it satisfies (16) for all $n \in \mathbb{N}$; a *continuum* of $\mathbb{R}^K$ is a connected compact subset of $\mathbb{R}^K$; we call $x \in \Xi$ *isolated* if

$$\exists \alpha > 0,\ \forall y \in \Xi \setminus \{x\},\ \|y - x\| \geqslant \alpha;$$

and Ξ is said to be *discrete* if all its points are isolated.

Theorem 1: Assume that conditions C1, C2, C3 and C5 are satisfied and that either C4 or C4$'$ holds. Let $(x^{(n)})_n$ be any sequence generated by (16).

(I) $(x^{(n)})_n$ has convergent subsequences and all its accumulation points are in Ξ.

(II) $(V(x^{(n)}))_n$ is strictly decreasing if $x^{(0)} \notin \Xi$, and there exists $\xi \in \Xi$ such that $\lim_{n\to\infty} V(x^{(n)}) = V(\xi)$.

(III) Either $(x^{(n)})_n$ converges or its accumulation points form a continuum.

(IV) Let ξ be an isolated critical point of V. If ξ is a minimum of V, then there exists an open neighborhood Ω of ξ such that $(x^{(n)})_n$ converges to ξ if $x^{(0)} \in \Omega$.

Theorem 2: Let $(x^{(n)})_n$ be any sequence generated by (16) under the same conditions as in Theorem 1.

(V) If V is strictly convex, then $(x^{(n)})_n$ converges to the global minimum of V.

(VI) If Ξ is discrete, then $(x^{(n)})_n$ converges to a critical point of V.

(VII) $\lim_{n\to\infty} \inf_{\xi\in\Xi} \|x^{(n)} - \xi\| \;=\; 0.$

Theorems 1 and 2 characterize the behavior of the algorithm in three mutually exclusive situations covering all possibilities: (i) when V is strictly convex, (ii) when V is not strictly convex and Ξ is discrete, and (iii) when Ξ is not discrete. The corresponding conclusions are the following.

(i) The cost function V is strictly convex if C5 holds and if the data penalty functions $\phi_1, \ldots, \phi_{K'}$ and the PFs $\theta_1, \ldots, \theta_{L'}$ are increasing and strictly convex. Therefore, from (V), convergence to the global minimum always occurs when the ϕ_k's and the θ_l's are strictly convex.

(ii) If V is not strictly convex, then, from (VI), the algorithm is guaranteed to converge to a critical point $\xi \in \Xi$ if Ξ is discrete. We cannot rigorously exclude the possibility that ξ is a maximum or a saddle point, but this is very unlikely since any isolated critical point that is a minimum is an attractor (by (IV)) whereas maxima and saddle points are unstable (by (II)).

(iii) If Ξ is not discrete, then, from (I) and (III), either $(x^{(n)})_n$ converges or Ξ contains a non-empty continuum. The algorithm also behaves well in the latter situation: by (II) and (VII), $(V(x^{(n)}))_n$ converges and $(x^{(n)})_n$ gets arbitrarily close to Ξ.

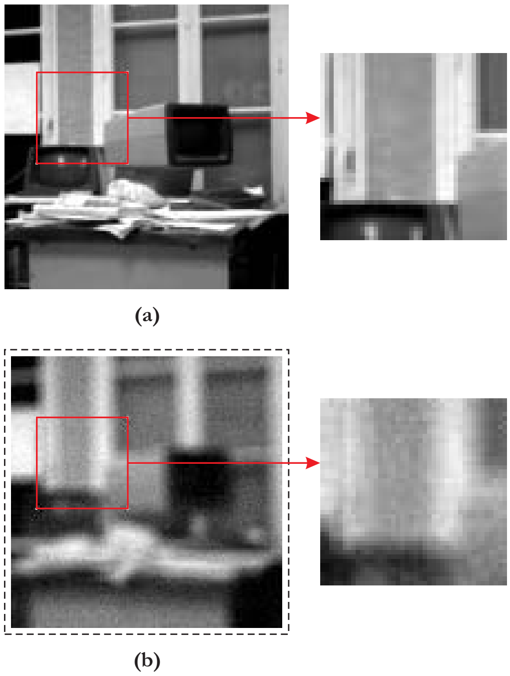


Fig. 2. (a) original image; (b) degraded observation: 7×7 uniform blur + 20 dB noise.

IV. Experiments

We consider the reconstruction of the 128×128 image $x^\star$ shown in Fig. 2(a) from the data d displayed in Fig. 2(b). The original image has intensity values ranging from 0 to 255 and the data were generated by blurring with a 7×7 uniform mask and adding white Gaussian noise at 20 dB SNR. (The standard deviation σ of the noise is defined via the decibel level of the signal-to-noise ratio: $(\mathrm{SNR})_{\mathrm{dB}} = 20\log_{10}(\sigma_{\mathrm{e}}/\sigma)$, where σ_{e} is the standard deviation of the exact data $\mathcal{H}(x^\star)$.) The metrics used to assess reconstruction quality are the mean-square error (MSE) and the improvement in SNR (ISNR); the ISNR associated with a solution x is defined by

$$(\mathrm{ISNR})_{\mathrm{dB}} \;=\; 20\log_{10}\left(\frac{\|\, x^\star|_{\mathsf{S}'} - d\,\|}{\|\, x^\star|_{\mathsf{S}'} - x|_{\mathsf{S}'}\,\|}\right), \tag{21}$$

where $y|_{\mathsf{S}'}$ denotes the restriction of y to the lattice S' supporting d (in image deblurring, the solution space $\mathbb{R}^K$ and the data space $\mathbb{R}^{K'}$ are supported by two rectangular lattices S and S' such that $\mathsf{S}' \subset \mathsf{S}$).

The estimates of the original image are obtained by minimizing the augmented cost function V given in (8) with the contour-line smoothing term defined by (9), using the algorithm presented in Section III. Since the components of the noise term are i.i.d. normal random variables, we take $\phi_k(t) \;=\; t^2$ for all k. However, our optimization algorithm works well for any set of data penalty functions satisfying C4, which includes in particular the standard smooth approximation to the identity $\mathrm{id}^{\delta}_{\mathbb{R}_+}$ (5) and the continuously differentiable

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:5, No:6, 2011

functions stemming from robust statistics [13]—hence our approach can deal with impulse or "salt-and-pepper" noise and with data corrupted by outliers.

Each operator $\mathcal{R}_l$ in the spatial-domain smoothing term Ψ computes a discrete approximation to the gradient at pixel l. The PFs ψ_l are the same for all l; we consider the convex smooth approximation to the identity, $\mathrm{id}^{\delta}_{\mathbb{R}_+}$, defined by (5) and the non-convex PF

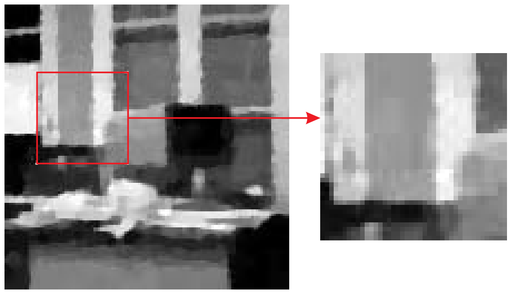
$$\psi^{\delta}_{\mathrm{HL}}(t) = \ln\bigl(1 + (t/\delta)^2\bigr). \tag{22}$$

The parameter δ associated with $\mathrm{id}^{\delta}_{\mathbb{R}_+}$ is set to 0.1 so that $\Psi(x)$ approximates the discrete TV of x. In the case of $\psi^{\delta}_{\mathrm{HL}}$, we take $\delta = 10$, and thus Ψ acts as a quadratic regularizer for gradient magnitudes up to 3–4 and as an edge detector for gradient magnitudes greater than 10.

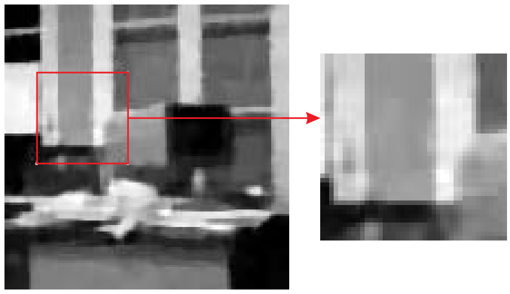
As regards the wavelet-domain smoothing term $\widetilde{\Psi}_{\mathrm{w}}$, we use the biorthogonal spline wavelet transform with two vanishing moments and two resolution levels. We take $\widetilde{\psi} = \mathrm{id}^{\delta}_{\mathbb{R}_+}$ with $\delta = 0.1$ and we set $(\mu_0, \mu_1) = (1, 2)$. The $\mathcal{D}^{\mathrm{v}}_j$'s compute differences between vertically adjacent coefficients in the horizontal detail images, and the $\mathcal{D}^{\mathrm{h}}_j$'s compute differences between horizontally adjacent coefficients in the vertical detail images.

Convex case: Fig. 3(a) shows the best reconstruction obtained when using $\psi_l = \mathrm{id}^{0.1}_{\mathbb{R}_+}$ without wavelet-domain smoothing. The corresponding value of the space-domain smoothing parameter λ is 1.0, the MSE is 208.0 and the ISNR is 5.70 dB. The edges are noisy, not to say visually unpleasant, and the staircase effect is clearly visible at the bottom of the image. By comparison, Fig. 3(b) displays the best solution achieved by adding the wavelet-domain smoothing term ($\widetilde{\lambda} = 0.8$). There are noticeable improvements: the image contours are smoother without penalizing object boundaries, and the "patchy" appearance due to the staircase effect is softened. Quantitatively, the MSE is 180.4 and the ISNR is 6.32dB. Fig. 4 displays the ISNR as a function of the wavelet-domain smoothing parameter $\widetilde{\lambda}$. Predictably, too large values of $\widetilde{\lambda}$ give over-smooth solutions. Yet, the values of $\widetilde{\lambda}$ that improve reconstruction quality cover a relatively large interval ($0 < \widetilde{\lambda} < 5.7$).

Non-convex case: Figs. 5 and 6 show the estimates obtained with $\psi_l = \psi^{10}_{\mathrm{HL}}$ for $(\lambda, \widetilde{\lambda}) = (12.6, 0)$ and $(\lambda, \widetilde{\lambda}) = (12.6, 2.2)$. The particular value $\lambda = 12.6$ was purposely selected to obtain a slightly under-regularized reconstruction in the absence of wavelet-domain smoothing (the optimal value is 25.2), and $\widetilde{\lambda} = 2.2$ is the optimal setting associated with $\lambda = 12.6$. The MSE and ISNR are 484.2 and 2.03 dB for the standard regularization approach, and 201.8 and 5.83 dB using our model. The improvements are even more striking than in the convex case: the edge artifacts are almost completely removed and the contour plots show substantial reduction of the staircase effect. As seen in Fig. 7, the evolution of the ISNR as a function of $\widetilde{\lambda}$ is reasonably smooth, which indicates that our optimization algorithm behaves well in the non-convex case. Improvements in reconstruction quality are observed for a wider range of values of $\widetilde{\lambda}$ than in the convex case (up to $\widetilde{\lambda} = 200$). Furthermore, even though λ is not optimal (the best reconstruction achieved without wavelet-domain smoothing has an MSE of 375.8 and an ISNR of 3.13 dB), our model outperforms optimal standard regularization over a significant range of values of $\widetilde{\lambda}$ ($0.12 < \widetilde{\lambda} < 60$).



(a)



(b)

Fig. 3. Reconstruction using the space-domain convex PF $\mathrm{id}^{0.1}_{\mathbb{R}_+}$: (a) without contour-line smoothing (MSE $= 208.0$, ISNR $= 5.70$dB); (b) with contour-line smoothing (MSE $= 180.4$, ISNR $= 6.32$ dB).

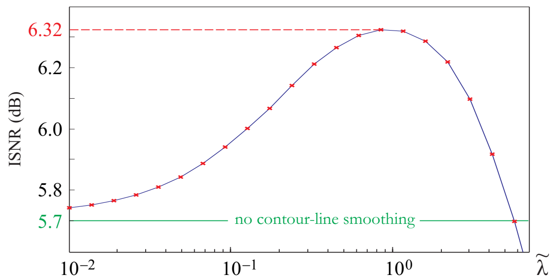


Fig. 4. ISNR as a function of the wavelet-domain smoothing parameter $\widetilde{\lambda}$ when using the space-domain convex PF $\mathrm{id}^{0.1}_{\mathbb{R}_+}$ (λ is set to its optimal value in terms of ISNR without contour-line smoothing).

V. Conclusion

We introduced the idea of "implicitly interacting" discontinuities by means of an additional penalty term operating on the detail images in the wavelet domain; the associated increase in optimization complexity is small compared to the use of an explicit line-processes. We also provided an efficient half-quadratic optimization algorithm which converges even when the faithfulness to the data is non-quadratic and the cost function is non-convex.

Compared to standard edge-preserving regularization, our model preserves boundary sharpness while smoothing contour lines. Aside from producing visually more pleasing reconstructions, this behavior is desirable for subsequent feature extraction and segmentation tasks; another benefit is the reduction of

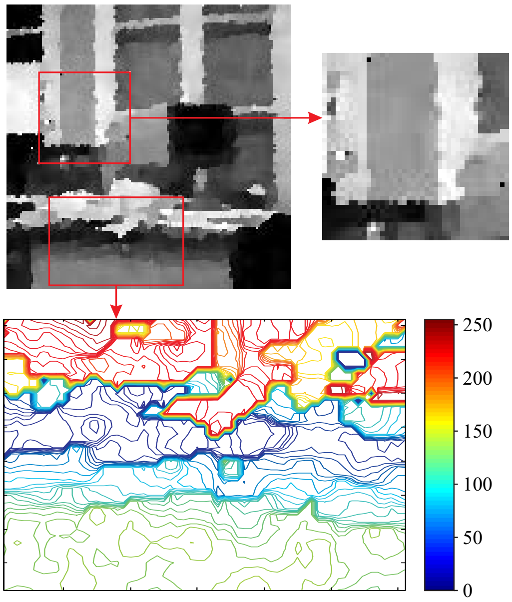


Fig. 5. Reconstruction using the space-domain non-convex PF ψ_{HL}^{10} without contour-line smoothing ($\mathrm{MSE} = 484.2$, $\mathrm{ISNR} = 2.03\,\mathrm{dB}$).

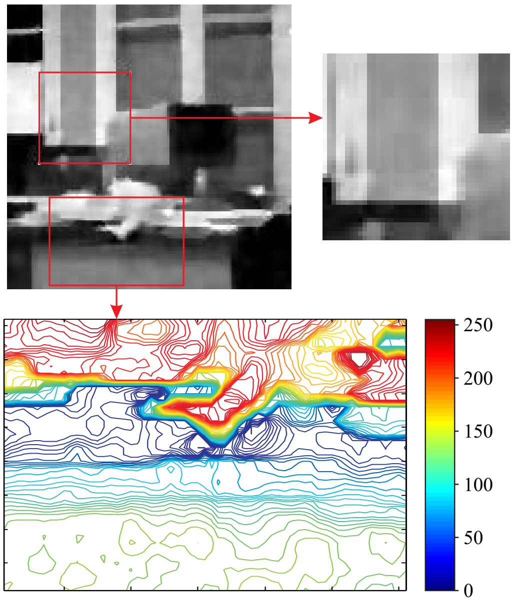


Fig. 6. Reconstruction using the space-domain non-convex PF ψ_{HL}^{10} and contour-line smoothing ($\mathrm{MSE} = 201.8$, $\mathrm{ISNR} = 5.83\,\mathrm{dB}$).

the staircase effect. The only drawback is the introduction of the additional hyper-parameter $\widetilde{\lambda}$ which controls the degree of smoothing in the wavelet domain. However, our experiments show that the range of values of $\widetilde{\lambda}$ that lead to an improvement in reconstruction quality is quite large, and preliminary results suggest that $\widetilde{\lambda}$ can be selected by tracking the value of the space-domain regularizer Ψ.

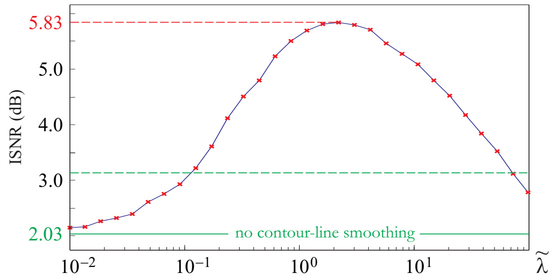


Fig. 7. ISNR as a function of the wavelet-domain smoothing parameter $\widetilde{\lambda}$ when using the space-domain non-convex PF ψ_{HL}^{10} (λ is set to half its optimal value and the highest ISNR achievable without contour-line smoothing is represented by the dotted green line).

References

[1] J. Marroquin, S. Mitter, and T. Poggio, "Probabilistic solution of ill-posed problems in computational vision," *J. Amer. Statist. Assoc.*, vol. 82, no. 397, pp. 76–89, 1987.

[2] G. Demoment, "Image reconstruction and restoration: overview of common estimation structures and problems," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, no. 12, pp. 2024–2036, 1989.

[3] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Trans. Image Process.*, vol. 6, no. 2, pp. 298–311, 1997.

[4] M. Nikolova, "Analysis of the recovery of edges in images and signals by minimizing nonconvex regularized least-squares," *Multiscale Model. Simul.*, vol. 4, no. 3, pp. 960–991, 2005.

[5] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 6, no. 6, pp. 721–741, 1984.

[6] L. Bedini, L. Benvenuti, E. Salerno, and A. Tonazzini, "A mixed-annealing algorithm for edge preserving image reconstruction using a limited number of projections," *Signal Process.*, vol. 32, no. 3, pp. 397–408, 1993.

[7] L. Blanc-Féraud, S. Teboul, G. Aubert, and M. Barlaud, "Nonlinear regularization using constrained edges in image reconstruction," in *Proc. IEEE Int. Conf. Image Processing*, vol. 2, (Lausanne, Switzerland), pp. 449–452, Sept. 1996.

[8] J. Idier, "Convex half-quadratic criteria and interacting auxiliary variables for image restoration," *IEEE Trans. Image Process.*, vol. 10, no. 7, pp. 1001–1009, 2001.

[9] D. Dobson and F. Santosa, "Recovery of blocky images from noisy and blurred data," *SIAM J. Appl. Math.*, vol. 56, no. 4, pp. 1181–1198, 1996.

[10] T. Chang, A. Marquina, and P. Mulet, "High-order total variation-based image restoration," *SIAM J. Sci. Comput.*, vol. 22, no. 2, pp. 503–516, 2000.

[11] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, no. 3, pp. 367–383, 1992.

[12] A. H. Delaney and Y. Bresler, "Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography," *IEEE Trans. Image Process.*, vol. 7, no. 2, pp. 204–221, 1998.

[13] M. J. Black and A. Rangarajan, "On the unification of line processes, outlier rejection, and robust statistics with applications in early vision," *Int. J. Comput. Vis.*, vol. 19, no. 1, pp. 57–91, 1996.