A l^1-Unified Variational Framework for Image Restoration

Julien Bect[1], Laure Blanc-Féraud[1], Gilles Aubert[2], and Antonin Chambolle[3]

[1] Projet Ariana, I3S/INRIA 2004 route des Lucioles, BP93
06902 Sophia Antipolis Cedex, France
blancf@sophia.inria.fr
http://www-sop.inria.fr/ariana/

[2] Laboratoire J-A. Dieudonné, Parc Valrose Univertsité de Nice-Sophia Antipolis
06108 Nice Cedex 2, France
gaubert@math.unice.fr
http://math.unice.fr/~gaubert/

[3] Ceremade, Univertsité Paris-Dauphine 75775 Paris Cedex 16, France
antonin@ceremade.dauphine.fr

Abstract. Among image restoration literature, there are mainly two kinds of approach. One is based on a process over image wavelet coefficients, as wavelet shrinkage for denoising. The other one is based on a process over image gradient. In order to get an edge-preserving regularization, one usually assume that the image belongs to the space of functions of Bounded Variation (BV). An energy is minimized, composed of an observation term and the Total Variation (TV) of the image.
Recent contributions try to mix both types of method. In this spirit, the goal of this paper is to define a unified-framework including together wavelet methods and energy minimization as TV. In fact, for denoising purpose, it is already shown that wavelet soft-thresholding is equivalent to choose the regularization term as the norm of the Besov space B_1^{11}. In the present work, this equivalence result is extended to the case of deconvolution problem. We propose a general functional to minimize, which includes the TV minimization, wavelet coefficients regularization, mixed (TV+wavelet) regularization or more general terms. Moreover we give a projection-based algorithm to compute the solution. The convergence of the algorithm is also stated. We show that the decomposition of an image over a dictionary of elementary shapes (atoms) is also included in the proposed framework. So we give a new algorithm to solve this difficult problem, known as Basis Pursuit. We also show numerical results of image deconvolution using TV, wavelets, or TV+wavelets regularization terms.

1 Introduction

1.1 Image Restoration

Restoring images from blurred or/and noisy data is an important task of image processing. In the important literature developed since twenty years, most

T. Pajdla and J. Matas (Eds.): ECCV 2004, LNCS 3024, pp. 1–13, 2004.

approaches are based on an energy minimization. Such energy contains mainly two terms: the first term models how the observed data is derived from the original data one would like to reconstruct; the second term contains a priori information on the regularity of this original data. At this point, two important families of criteria emerge. In the first family the regularity criterion is a semi-norm that is expressed in a “simple” way in terms of the wavelet coefficients of the image (usually a Besov norm). This leads to a restoration process that is performed through some processing of the wavelet coefficients, such as a wavelet shrinkage (for example see [5] in denoising, [10] in deconvolution, [8] in Radon transform inversion).
In the second family, the regularity criterion is a functional of the gradient of the image, so that the resolution of the problem amounts to solving some more or less complex PDE. In order to get an edge-preserving regularization, one usually assumes that the image belongs to the space of functions of Bounded Variation (BV) and the criterion which is minimized is the Total Variation (TV) of the image (see [12] for example).

Recent contributions try to mix both types of method [9,14,6]. In this spirit, the goal of this paper is to define a unified-framework including together wavelet, TV, or a more general semi-norm. In fact, as it is shown in [2] for denoising and compression purposes, wavelet soft-thresholding is equivalent to choose the regularization term as the norm of the Besov space B_1^{11}. In the present work, this equivalence result is extended to the case of deconvolution problem. The proposed framework allows to include the TV minimization, mixed (TV+wavelet) regularization or more general terms. Moreover we give a projection-based algorithm to compute the solution in the more general case. The convergence of algorithm is also stated.

Image restoration can be considered as the minimization of a functional written as

$$\frac{1}{2\lambda}\,\|g - Au\|_{X_1}^2 + |u|_Y^s \tag{1}$$

A is a linear operator which can model the degradation during the observation of the object u:

$$\begin{aligned} X &\longrightarrow X_1 \\ u &\longmapsto g = Au + \eta \end{aligned} \tag{2}$$

X is the space describing the objects to be reconstructed and X_1 the space of observations. η is the acquisition noise. Typically $X = X_1 = L^2$ or $X = X_1$ a finite-dimensional space. As in [2], $|u|_Y$ is a norm or a semi-norm in a smoothness space Y. Standard example is $Y = L^2, s = 2$ defining quadratic regularization as proposed by Tikhonov [15]. Now, if Y is the BV space and $s = 1$, the solution is the one such that Au best approximates g (in the sense of the norm $\| \, \|_{X_1}$), with minimal Total Variation [12]. This general functional includes also wavelet shrinkage denoising/deconvolution methods by considering $A = I$ (where I is the identity operator) or A is the Point Spread Function of the transfert function of the optics and Y is the Besov space B_1^{11} and $s = 1$ [2]. Notice that if A defines a decomposition over a dictionary of possible atoms from which the signal u is built, (for example wavelet packets for textures, curvelets or bandlets for edges, and so on), then solving (1) corresponds exactly to the Basis Pursuit DeNoising algorithm (BPDN) by Chen, Donoho and Saunders [3].

1.2 Problem Statement

In this paper, we study the minimization of a functional of the form

$$\frac{1}{2\lambda}\|g - Au\|_{X_1}^2 + J(u) \tag{3}$$

where $J : X \to \mathbb{R}\cup+\{\infty\}$ is a semi-norm on X. For sake of simplicity, we assume in the whole paper that u is a discrete image that is to say $X = X_1 = \mathbb{R}^{N\times N}$ and the symbol $\|.\|$ will denote any Hilbertian norm.

In order to minimize the functional (3), we describe a projection-based algorithm which extends the one proposed by Chambolle for the denoising case $(A = I)$ with TV regularization [1].

The convergence of the algorithm is proved. This gives a new algorithm to solve several kinds of image processing: image deconvolution with TV regularization, with wavelet shrinkage, or with both kind of regularization; BPDN problem as decribed by Donoho in [3]. During the review process of the paper, our attention was drawn by S. Mallat to the independent works [7,4] which derive, by different approaches, essentially the same iterative algorithm as the one described in this paper. In [4], a strong convergence of the iterative algorithm is shown in infinite dimension. One difference is that our algorithm includes TV or mixed (TV+wavelet) regularization which seems not to be the case in [7,4].

In section 2, we recall some basic tools in convex analysis. The main contribution of this paper is detailed in section 3 where the minimization algorithm is given for the general functional (3). In section 4, we show that several standard methods in image restoration are special cases of the unified energy (3), and numerical results are given for deconvolution with TV plus wavelet regularization.

1.3 Notations

Let us fix some notations. A discrete image will be denoted by $u_{i,j}$, $i, j = 1 \ldots N$. In order to define the TV of the discrete image u, we introduce the gradient $\nabla : X \to X \times X$ defined by:

$$(\nabla u)_{i,j}^1 = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } i < N \\ 0 & \text{if } i = N \end{cases} \quad \text{and} \quad (\nabla u)_{i,j}^2 = \begin{cases} u_{i,j+1} - u_{i,j} & \text{if } j < N \\ 0 & \text{if } j = N \end{cases}$$

We also introduce a discrete version of the divergence operator defined, by analogy with the continuous case, by $\mathrm{div} = -\nabla^*$ where ∇^* is the adjoint of ∇.

We have

$$(\operatorname{div}(p))_{i,j} = \begin{cases} p^1_{i,j} - p^1_{i-1,j} & \text{if } 1 < i < N \\ p^1_{i,j} & \text{if i=1} \\ -p^1_{i-1,j} & \text{if i=N} \end{cases} + \begin{cases} p^2_{i,j} - p^2_{i,j-1} & \text{if } 1 < j < N \\ p^2_{i,j} & \text{if j=1} \\ -p^2_{i,j-1} & \text{if j=N} \end{cases} \quad (4)$$

The discrete TV denoted J_{TV} is defined as the l^1-norm of the vector ∇u by

$$J_{TV}(u) = \|\nabla u\|_1 = \sum_{i,j=1}^{N} \sqrt{\left\{(\nabla u)^1_{i,j}\right\}^2 + \left\{(\nabla u)^2_{i,j}\right\}^2}. \quad (5)$$

2 Some Tools of Convex Analysis

We recall in this section some usual tools in convex analysis which are used to build our algorithm. We refer the reader to Rockafellar [11] for a more complete introduction of convex analysis.

Definition 1 (Legendre-Fenchel Conjugate). *Let ϕ be an application $X \to \mathbb{R} \cup \{+\infty\}$. We assume that $\phi \not\equiv +\infty$. The conjugate function of ϕ is defined as $\phi^* : X \to \mathbb{R} \cup \{+\infty\}$ by:*

$$\phi^*(s) \; = \; \sup_{x \in X} \, \{ \, \langle \, s, \, x \, \rangle \, - \, \phi(x) \, \} \quad (6)$$

ϕ^* is convex and lower semi-continuous (lsc).

Definition 2 (Indicator function, support function).
Let $K \subset X$ be a non empty closed convex subset of X. The indicator function of K, called χ_K, is defined by:

$$\chi_K(x) \; = \; \begin{cases} 0 & \textit{if } x \in K \\ +\infty & \textit{otherwise} \end{cases} \quad (7)$$

We call support function of K the function denoted δ_K, defined by:

$$\delta_K(s) \; = \; \sup_{x \in K} \, \langle \, s, \, x \, \rangle \quad (8)$$

The link between χ_K and δ_K is given by the following results

Theorem 1. *Let $K \subset X$ a non empty convex closed subset of X. Then the functions χ_K and δ_K are convex, lsc and mutually conjugate.*

Theorem 2. *All support functions δ_K, associated to a non empty convex closed subset K, is convex and one-homogeneous (e.g. $\forall t > 0, \forall x \in X$, $\delta_K(tx) = t\delta_K(x)$) and lsc. Conversely, each function $\phi \not\equiv +\infty$ convex, one-homogeneous and lsc is the support function of a closed convex set K_ϕ defined by:*

$$K_\phi \; = \; \{ \, s \in X, \quad \forall \, x \in X, \, \langle \, s, \, x \, \rangle \, \leq \phi(x) \, \} \, . \quad (9)$$

For example, in (3), we have supposed that $J(u)$ is a semi-norm. Therefore, it is a convex one-homogeneous and lsc function. So $J(u)$ is the support function of a closed convex set K_J defined by (9). If $J(u)$ is the discrete TV semi-norm given by (5), then

$$K_{TV} = \{ \operatorname{div}(p), \quad p \in X \times X,\ |p_{i,j}| \leq 1\ \forall i,j \}\,. \tag{10}$$

We now introduce the notion of sub-differential of a function which generalizes the differential for convex functions.

Definition 3 (Sub-differential). *Let ϕ a convex function. We define the sub-differential $\partial\phi(x)$ of ϕ in $x \in X$ by:*

$$s \in \partial\phi(x) \qquad \Longleftrightarrow \qquad \forall\, x' \in X, \quad \phi(x') \;\geq\; \phi(x) + \langle\, s,\, x' - x\,\rangle \tag{11}$$

Note that if x is such that $\phi(x) < \infty$ and if ϕ is differentiable in x, then:

$$\partial\phi(x) \;=\; \{\, \nabla\phi(x)\,\}\,. \tag{12}$$

3 Algorithm and Convergence Result

This section is devoted to the main contribution of this paper, namely the description and the convergence of our algorithm for numerically solving the minimization problem (3). Before doing that, in order to justify our algorithm, we need some preliminary results.

3.1 Preliminary Results

Theorem 3. *Let $B : X \to X$ be a linear self-adjoint and positive operator satisfying $\|B\| < 1$. Then*

$$\forall u \in X, \quad \langle\, Bu,\, u\,\rangle \;=\; \min_{w \in X} \left\{ \|u - w\|^2 + \langle\, Cw,\, w\,\rangle \right\} \tag{13}$$

where $C = B\,(I - B)^{-1}$. Moreover, the minimum is reached at a unique point w_u which verifies:

$$w_u \;=\; (I + C)^{-1}\,(u) \;=\; (I - B)\,(u)\,. \tag{14}$$

Let us recall that the functional we want to minimize is given by

$$\frac{1}{2\lambda}\,\|g - Au\|^2 + J(u)$$

Let $\mu > 0$ be such that

$$\mu \|A^* A\| < 1 \tag{15}$$

and

$$B = \mu A^* A \tag{16}$$

B is a self-adjoint positive operator. From hypothesis(15), μ is such that $\|B\| < 1$. Thanks to Theorem 3, we will be able to write the data term of (3), $(\frac{1}{2\lambda}\|g - Au\|^2)$, as the result of a new minimization problem, w.r.t an auxiliary variable w. We have

$$\|Au\|^2 = \langle A^* Au, u \rangle \tag{17}$$

$$= \frac{1}{\mu} \langle Bu, u \rangle \tag{18}$$

$$= \frac{1}{\mu} \min_{w \in X} \left\{ \|u - w\|^2 + \langle Cw, w \rangle \right\} \tag{19}$$

with $C = B(I - B)^{-1}$. Therefore

$$\frac{1}{2\lambda} \|g - Au\|^2 = \min_{w \in X} H(u, w) \tag{20}$$

where H is the convex differentiable function defined by:

$$H(u, w) = \frac{1}{2\lambda\mu} \left(\|u - w\|^2 + \langle Cw, w \rangle \right) + \frac{1}{2\lambda} \left(\|g\|^2 - 2 \langle Au, g \rangle \right). \tag{21}$$

Let us denote

$$\Psi_1 = I - B = I - \mu A^* A \tag{22}$$

From relation (14), w minimizes $H(u, .)$ if and only if $w = \Psi_1 u$. Let us now consider the function F defined by:

$$F(u, w) = H(u, w) + J(u) \tag{23}$$

F is a convex continuous function, and we deduce from the previous preliminary results, the following proposition:

Proposition 1. *w minimizes $F(u, .)$ defined in (21) and (23) if and only if $w = \Psi_1 u$ where $\Psi_1 = I - \mu A^* A$ and we have:*

$$\forall w \neq \Psi_1 u, \quad F(u, \Psi_1 u) < F(u, w) \tag{24}$$

Let us now show that computing the global minimizer of F reduces to minimize each of its partial functions $F(., w)$ and $F(u, .)$. This is a non trivial result even in the case of a strictly convex function (consider for instance $f(x, y) = (x^2 + y^2)/2 + |x - y|$ at $x = y = 1/2$).

Proposition 2. *(u, w) minimizes F if and only if:*

$$\begin{cases} u \text{ minimizes } F(\,.\,, w) \\ w \text{ minimizes } F(\,u, .\,) \end{cases} \tag{25}$$

Proof. Since F is the sum of two convex continuous functions,

$$\partial F(u,w) \;=\; \partial H(u,w) \;+\; \partial J(u,w). \tag{26}$$

As H is differentiable and J does not depend on w, we deduce:

$$\partial F(u,w) \;=\; (\,\nabla_u H(u,w)\,,\, \nabla_w H(u,w)\,) \;+\; \partial J(u) \times \{0\} \tag{27}$$

So

$$0 \in \partial F(u,w) \quad \Leftrightarrow \quad \begin{cases} 0 \in \nabla_u H(u,w) \;+\; \partial J(u) \\ 0 = \nabla_w H(u,w) \end{cases} \tag{28}$$

which is exactly what we want to show.

The last result we need in order to derive our algorithm is the following:

Proposition 3. *Let us denote by $\Psi_2 : X \to X$ the application defined by:*

$$\Psi_2(w) \;=\; (I - \Pi_{\lambda\mu K_J})\,(w \;+\; \mu A^* g)\,. \tag{29}$$

Here $\Pi_{\lambda\mu K_J}(w)$ stands for the orthogonal projection of w on the convex set $\lambda\mu K_J$, where K_J is the convex set associated to $J(u)$ (see (9)). Then u minimizes $F(\,.\,,w)$ if and only if $u = \Psi_2(w)$.

The expression (29) is found by computing the dual problem of $\min_u F(u,w)$, for fixed w (see [1]).

3.2 The Algorithm for the Minimization of the Unified Functional

We are now able to describe the algorithm we propose to minimize the unified functional (3). Based on results given in Propositions 1, 2 and 3, we propose the following iterative algorithm to minimize (3)

$$w_n \;=\; (I - \mu A^* A)\,(u_n) \tag{30}$$

$$u_{n+1} \;=\; (I - \Pi_{\lambda\mu K_J})\,(w_n + \mu A^* g) \tag{31}$$

By a change of notation, using $v_n = w_n + \mu A^* g$, it results:

$$v_n \;=\; u_n \;+\; \mu\, A^*\,(g - A u_n) \tag{32}$$

$$u_{n+1} \;=\; (I - \Pi_{\lambda\mu K_J})\,(v_n) \tag{33}$$

In practice, we will use the algorithm (32)–(33) rather than the writting (30)–(31) and we use the numerical algorithm (35)–(36) described in section 3.3 to compute $\Pi_{\lambda\mu K_J}$.

The first equation of this algorithm is a fixed-step descent algorithm, considering only the minimization of the data term $\|g - Au\|^2$. The step is fixed by the parameter μ. The second equation corresponds to a denoising step over the current estimates v_n. Remark that the parameter considered in the denoising step is $\lambda\mu$ rather than λ as it should be suggested looking at the functional (3). We can also observe that in the case where A^*A is invertible then (32)–(33) correspond to a contraction with a ratio $1-\mu\lambda_0$, where λ_0 is the smallest eigenvalue of A^*A. In the case where A^*A is not invertible, then the transformation (32)–(33) is 1–Lipschitz. In either situations the following theorem holds.

Theorem 4 (Convergence of the algorithm). *Let $\mu > 0$ and assume*

$$\mu \, \|A^*A\| \; < \; 1 \tag{34}$$

Then the algorithm (32)–(33) *converges to a global minimizer of (3).*

Before ending this section, let us remark that we always have the existence of a minimizer of the functional (3) in the discrete setting. However the difficult point for the convergence proof comes from the fact that the minimum is non necessarily unique.

3.3 Projection Algorithm of Chambolle

We give in this section the numerical algorithm to compute a projection $\Pi_{\lambda K_J}$, in the case of a regularizing term expressed as $J(u) = \|Qu\|_1 = \sum_\theta |(Qu)_\theta|$, where Q is a linear operator $Q : X \to \Theta$ and Θ is a product space (see section 4.3 for more details on the notations). The projection onto the convex closed set λK_J, where K_J is associated to $\|Qu\|_1$ can be numerically computed by a fixed point method, based on results in [1]. We build recursively a sequence in Θ of vectors $p_n = (p_{n,\theta})_\theta$ in the following way: we choose $p_0 \in B_\Theta = \{p \in \Theta \;:\; |p_\theta| \le 1 \;\forall\theta\}$ and for each $n \ge 0$ we let

$$q_n = Q(Q^* p_n - \frac{g}{\lambda}) \tag{35}$$

and for each θ

$$p_{n+1,\theta} = \frac{p_{n,\theta} - \tau(q_{n,\theta})}{1 + \tau|q_{n,\theta}|} \tag{36}$$

We have a sufficient condition ensuring the convergence of the algorithm:

Theorem 5. *Assume that the parameter τ in (36) verifies $\tau \le \frac{1}{\kappa^2}$ where $\kappa = \|Q^*\|$. Then for all initial condition $p_0 \in B_\Theta$, the algorithm* (35)–(36) *is such as:*

$$\lambda \, Q^* p_n \;\longrightarrow\; \lambda \, Q^* \hat{p} \;=\; \Pi_{\lambda K_J}(g) \tag{37}$$

4 Applications

For TV regularization, one just needs to apply the algorithm (32)–(33) with $K = K_{TV}$, so that the projection algorithm is given by (35)–(36) with $\Theta = X \times X$, $Q = \nabla$, $Q^* = -\mathrm{div}$ (as described in [1]). Let us now look at regularization in the wavelet domain.

4.1 Wavelet Shrinkage

Let us consider the case where $J(u)$ is the norm in the Besov space B_1^{11}. In [2], it is shown that this norm is equivalent to the norm of the sequence of wavelet coefficients $c^u_{j,k,\psi}$:

$$\|u\|_{B_1^{11}} = \sum_{j,k,\psi} |c^u_{j,k,\psi}| \tag{38}$$

$\psi \in \{\psi^{(1)}, \psi^{(2)}\psi^{(3)}\} = \Psi$ defines bi-dimensional wavelets from a one-dimensional wavelet and a one-dimensional scaling function as usual. The set of functions $\{\psi_{j,k}(x) = 2^k\psi(2^k x - j)\}_{\psi\in\Psi, k\in Z, j\in Z^2}$ forms an orthogonal bases for $L^2(\mathbb{R}^2)$. Then, for $f \in L^2(\mathbb{R}^2)$, we have

$$f = \sum_{j,k,\psi} c^f_{j,k,\psi}\psi_{j,k} \tag{39}$$

For sake of simplicity, the range of the indexes is omitted: we work with discrete functions with bounded definition domain. Assume that our purpose is image deconvolution that is to say A is a convolution operator representing the transfert function of the optics. Then if we want to deconvolve the observed image g with a wavelet regularization term, we have to minimize an energy of the form

$$\frac{1}{2\lambda}\|g - Au\|^2 + \|u\|_{B_1^{11}}. \tag{40}$$

We know (see for example [2]), that if $A = I$, minimizing (40) is equivalent to a soft-thresholding algorithm. Let us now see what happens with algorithm (32)–(33) and a general operator A. The convex set K_1 associated to the norm in B_1^{11} is defined by (see (9))

$$K_1 = \left\{ s \in X \,/\, \forall x \in X, < s, x > \leq \|x\|_{B_1^{11}} \right\}. \tag{41}$$

We easily deduce that

$$K_1 = \left\{ s \in X \,/\, \forall j, k, \psi, \; |c^s_{j,k,\psi}| \leq 1 \right\}. \tag{42}$$

where c^s are the wavelet coefficients of s

Therefore equation (33) is a denoising step by soft-thresholding with threshold $\lambda\mu$. Then the algorithm iteratively computes a step of steepest gradient descent only for the deconvolution and then a denoising step by soft-thresholding. This algorithm is very easy to implement.

4.2 Basis Pursuit DeNoising (BPDN)

Representing a signal in terms of few high coefficients of a dictionary and a lot of vanishing coefficients allows representation of an image ensuring better performances of shrinkage methods or other restoration methods. The problem is to decompose a signal over a possibly large dictionary rather than one orthogonal basis. The dictionary should contain all possible atoms which can be used to represent any images. For example we can use in the dictionary DCT, DST, biorthogonal wavelets, wavelet packets, curvelets, and so on. Searching this representation is an ill-posed problem, since such a decomposition is non unique. In the Basis Pursuit DeNoising algorithm (BPDN) [3] the authors propose the following regularizing functional

$$\inf_{\alpha} \frac{1}{2\lambda} \|g - \Phi\alpha\|^2 + \|\alpha\|_1 \tag{43}$$

The function g is the signal to be decomposed, α the unknown coefficients and Φ the operator $\Phi : \alpha \longmapsto y = \sum_i \alpha_i \phi_i$ where ϕ_i are the elements of the dictionary.

The minimization (43) can be performed by using algorithm (32)–(33). Since the regularization is a l^1-norm, the step (33) is simply a soft-thresholding.

In [3], is proposed the algorithm IP (Interior Point) to solve (43). This algorithm is slow. A faster algorithm called BCR (Block Coordinate Relaxation) has been proposed in [13]. As algorithm (32)–(33), BCR is based on a soft-thresholding of the coefficients. In BCR, it is assumed that the dictionary is composed of a union of orthogonal bases. Our algorithm is more general since it can be applied by using any dictionary.

We have compared these three algorithms on some 1D-signals. The IP algorithm is available on the web (*http://www-stat.stanford.edu/atomizer*). We have chosen the same bases (wavelet transform, DCT and DST) for the three algorithms. On a 1D signal of 4096 samples, it appears that the convergence is much faster for algorithm (32)–(33) than the IP algorithm, and a little bit smaller than the BCR one. We may loose in time what we gain in generality. Of course these are very few results and much more experiments must be conducted for the comparison.

4.3 l^1-Regularization

In this section, we show that our algorithm can be applied to a general class of semi-norm $J(u)$ which is relevant in real problems. We will consider what we call l^1-regularization, which consists in the minimization of the following functional

$$\frac{1}{2\lambda} \|g - Au\|^2 + \|Qu\|_1. \tag{44}$$

Q is a linear application $Q : X \to \Theta$, where Θ is the product space defined as:

$$\Theta = \prod_{1 \leq \theta \leq r} \mathbb{R}^{n_\theta} \tag{45}$$

endowed with the norm

$$p \in \Theta \longmapsto \|p\|_1 = \sum_{1 \leq \theta \leq r} |p|_\theta \tag{46}$$

where $|\cdot|$ is the Euclidean norm on $\mathbb{R}^{n_\theta}$ and $p = (p_\theta)_{1\leq\theta\leq r}$, $p_\theta \in \mathbb{R}^{n_\theta}$.

$\|Qu\|_1$ is a semi-norm over X and is a norm when Q is injective.
For example, if Q is a wavelet transform, $(Qu)_\theta$ are scalar coefficients and then

$$\|Qu\|_1 \;=\; \sum_\theta |\,(Qu)_\theta\,| \tag{47}$$

is a norm if the sum runs over all coefficients of the wavelet transform or a semi-norm otherwise.
In the TV case we have

$$\|Qu\|_1 \;=\; \|\nabla u\|_1 \tag{48}$$

This general framework also includes a regularization composed of a sum of a TV term and a wavelet term. For such a regularization, we will set $\Theta = X^2 \times X$, and $Q : X \to \Theta$ is defined as:

$$Q: \quad X \longrightarrow X^2 \times X \tag{49}$$

$$u \longmapsto \begin{pmatrix} \gamma\,\nabla u \\ (1-\gamma)\,Wu \end{pmatrix} \tag{50}$$

where W stands for an orthonormal wavelet transform. We use the norm $\|\cdot\|_1$ such that: $\forall\,(p^1, p^2) \in X^2$, $\forall\, w \in X$,

$$\|(p,w)\|_1 = \|p\|_1 \,+\, \|w\|_1 \tag{51}$$

$$= \sum_{1\leq i,j\leq N} \sqrt{\left(p^1_{i,j}\right)^2 + \left(p^2_{i,j}\right)^2} \;+\; \sum_{1\leq i,j\leq N} |w_{i,j}| \tag{52}$$

and the global regularization functional is

$$J_\gamma(u) \;=\; \gamma\,\|\nabla u\|_1 \,+\, (1-\gamma)\,\|Wu\|_1 \;=\; \|Qu\|_1 \tag{53}$$

Note that (53) defines a family of functional J_γ, which goes continuously from l^1-norm on the wavelet coefficient to the Total Varation as γ goes from 0 to 1. We show restoration results for deconvolution by using this functional for three values of γ. The Lena image has been blurred by a PSF (Point Spread Function) corresponding to a synthetic aperture optical system, with vanishing coefficients in the medium frequencies as well as in the high frequencies. A Gaussian white noise has been added with standard deviation $\sigma = 0.05$ (for u values in $[0,1]$). This deconvolution problem is very difficult because of vanishing medium frequencies of the degradation, and a large amount of noise. We retrieve for $\gamma = 0$ and $\gamma = 1$ the specific drawbacks of wavelets and TV restoration respectively: blur and bad edges for the wavelets, loose of textures for TV. The value $\gamma = 0.2$ has been chosen by hand and gives a good compromise. The choice of this parameter is an open problem. The regularizing parameter λ is estimated following the ideas of [1].

Original image — blurred image

Observations (blurred image + noise, $\sigma = 0.05$) — Deconvolution with $\gamma = 0$

Deconvolution with $\gamma = 0.2$ — Deconvolution with $\gamma = 1$

Fig. 1. Results of the algorithm on a deconvolution problem

5 Conclusion

We have presented a general functional unifying several approaches of image restoration. A convergent and easy to implement algorithm has been proposed for the minimization of this functional. For a good evaluation of our algorithm in terms of quality and rapidity for several applications, much more results will be conducted in each specific application as deconvolution or BPDN.

References

1. A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1-2):89–97, 2004. Special Issue on Mathematics and Image Analysis.
2. A. Chambolle, R. DeVore, N. Lee, and B. Lucier. Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage. *IEEE Trans. Image Proc.*, 7(3):319–333, 1998.
3. Scott Shaobing Chen, David L. Donoho, and Michael A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1999.
4. Ingrid Daubechies, Michel Defrise, and Christine De Mol. An iterative algorithm for linear inverse problems with a sparsity constraint. Technical report, http://arxiv.org/abs/math.FA/0307152, 2003.
5. David L. Donoho and Iain M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.
6. Sylvain Durand and Jacques Froment. Reconstruction of wavelet coefficients using total variation minimization. *SIAM Journal of Scientific Computing*, 24(5):1754–1767, 2003.
7. Mario A.T. Figueiredo and Robert D. Nowak. An EM algorithm for wavelet-based image restoration. *IEEE Transactions on Image Processing*, 12(8), august 2003.
8. Namyong Lee and Bradley J. Lucier. Wavelet methods for inverting the radon transform with noisy data. *IEEE Transactions on Image Processing*, 10:79–94, 2001.
9. Fran cois Malgouyres. Minimizing the total variation under a general convex constraint for image restoration. *IEEE transactions on image processing*, 11(12):1450–1456, 2002.
10. Stéphane Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, 1998.
11. R.T. Rockafellar. *Convex Analysis*. Princeton University press, 1972.
12. L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithm. *Physica D*, 60:259–268, 1992.
13. S. Sardy, A. Bruce, and P. Tseng. Block coordinate relaxation methods for nonparametric signal denoising with wavelet dictionaries. Technical report, Seattle, WA, 1998.
14. Jean-Luc Starck, Fionn Murtagh, and Albert Bijaoui. Multiresolution support applied to image filtering and restoration. *Graphical Models and Image Processing*, 57(5):420–431, 1995.
15. A.N. Tikhonov and V.Y. Arsenin. *Solutions of ill-posed problems*. Winston and Wiley, 1977.