

Proximal Algorithms for Multicomponent Image Recovery Problems

L.M. Briceño-Arias · P.L. Combettes · J.-C. Pesquet · N. Pustelnik

Published online: 21 December 2010
© Springer Science+Business Media, LLC 2010

Abstract In recent years, proximal splitting algorithms have been applied to various monocomponent signal and image recovery problems. In this paper, we address the case of multicomponent problems. We first provide closed form expressions for several important multicomponent proximity operators and then derive extensions of existing proximal algorithms to the multicomponent setting. These results are applied to stereoscopic image recovery, multispectral image denoising, and image decomposition into texture and geometry components.

Keywords Convex minimization · Image recovery · Inverse problems · Multicomponent images · Multichannel images · Multispectral images · Proximal algorithm · Sparsity · Stereoscopy · Wavelets

This work was supported by the Agence Nationale de la Recherche under grants ANR-08-BLAN-0294-02 and ANR-09-EMER-004-03.

L.M. Briceño-Arias · P.L. Combettes (✉)
Laboratoire Jacques-Louis Lions—CNRS UMR 7598,
UPMC Université Paris 06, 75005 Paris, France
e-mail: plc@math.jussieu.fr

L.M. Briceño-Arias
e-mail: lbriceno@math.jussieu.fr

L.M. Briceño-Arias
Équipe Combinatoire et Optimisation—CNRS FRE 3232,
UPMC Université Paris 06, 75005 Paris, France

J.-C. Pesquet · N. Pustelnik
Laboratoire d'Informatique Gaspard Monge—CNRS UMR 8049,
Université Paris-Est, 77454 Marne la Vallée Cedex 2, France

J.-C. Pesquet
e-mail: jean-christophe.pesquet@univ-paris-est.fr

N. Pustelnik
e-mail: nelly.pustelnik@univ-paris-est.fr

1 Problem Statement

In this paper, we consider signal and image recovery problems in which the ideal solution is represented by m components x_1, \dots, x_m lying, respectively, in real Hilbert spaces $\mathcal{H}_1, \dots, \mathcal{H}_m$. Such problems arise in many areas ranging from color and hyperspectral imaging to multichannel signal processing and geometry/texture image decomposition [2, 4–7, 12, 23, 25, 29, 30, 40, 44, 46]. Oftentimes, multicomponent signal/image processing tasks can be formulated as variational problems of the form

$$\underset{x_1 \in \mathcal{H}_1, \dots, x_m \in \mathcal{H}_m}{\text{minimize}} \quad \Phi(x_1, \dots, x_m), \quad (1)$$

where Φ is a convex function modeling the available information on the m components, their interactions, and, possibly, the data acquisition process.

The abstract convex minimization problem (1) is usually too generic to be solved directly and it must be formulated in a more structured fashion to be amenable to efficient numerical solution. To this end, Φ can be decomposed as a sum of p functions that can be handled individually more easily. This leads to the following model, which will be the focus of the paper.

Problem 1 Let $(\mathcal{H}_i)_{1 \leq i \leq m}$ be real Hilbert spaces, and let $(f_k)_{1 \leq k \leq p}$ be proper lower semicontinuous convex functions from the direct Hilbert sum $\mathcal{H}_1 \oplus \dots \oplus \mathcal{H}_m$ to $]-\infty, +\infty]$. The problem is to

$$\underset{x_1 \in \mathcal{H}_1, \dots, x_m \in \mathcal{H}_m}{\text{minimize}} \quad \sum_{k=1}^p f_k(x_1, \dots, x_m), \quad (2)$$

under the assumption that solutions exist.

In the case of univariate ($m = 1$) signal processing problems, proximal methods have been successfully used to solve (2); see [17, 19, 21] for basic work, and [20] and the references therein for a variety of applications. It is therefore natural to ask whether these methods can be extended to the multivariate case. Initial work in this direction was carried out in [10] in the special instance when $m = 2$, f_1 is a separable sum (i.e., $f_1 : (x_i)_{1 \leq i \leq m} \mapsto \sum_{i=1}^m \varphi_i(x_i)$), and f_2 is differentiable on $\mathcal{H}_1 \oplus \dots \oplus \mathcal{H}_m$ with a Lipschitz continuous gradient (this setting also covers formulations found in [4–6, 21, 26, 27, 41, 42, 46]). The objective of our paper is to address the general case and to present several proximal algorithms with guaranteed convergence to a solution to Problem 1 under suitable assumptions.

The paper is organized as follows. In Sect. 2, the main notation used in the paper is introduced. Proximity operators will be an essential ingredient in the multicomponent algorithms proposed in the paper. They are briefly reviewed in Sect. 3, where we also provide new results concerning multicomponent proximity operators. In Sect. 4, we describe proximal splitting algorithms which are pertinent for solving Problem 1. Finally, in Sect. 5, we illustrate the effectiveness of the proposed algorithms in three multicomponent imaging examples.

2 Notation

Throughout, \mathcal{H} , \mathcal{G} , and $(\mathcal{H}_i)_{1 \leq i \leq m}$ are real Hilbert spaces. For convenience, their scalar products are all denoted by $\langle \cdot | \cdot \rangle$, the associated norms by $\| \cdot \|$, and their identity operators are all denoted by Id . It will be convenient to denote by $\mathbf{x} = (x_i)_{1 \leq i \leq m}$ a generic element in $\mathcal{H}_1 \times \dots \times \mathcal{H}_m$, and by \mathcal{H} the direct Hilbert sum $\mathcal{H}_1 \oplus \dots \oplus \mathcal{H}_m$, i.e., the product space $\mathcal{H}_1 \times \dots \times \mathcal{H}_m$ equipped with the usual vector space structure and the scalar product

$$(\mathbf{x}, \mathbf{y}) \mapsto \sum_{i=1}^m \langle x_i | y_i \rangle. \tag{3}$$

The space of bounded linear operators from \mathcal{H} to \mathcal{G} is denoted by $\mathcal{B}(\mathcal{H}, \mathcal{G})$. Moreover, $\Gamma_0(\mathcal{H})$ denotes the class of lower semicontinuous convex functions $\varphi : \mathcal{H} \rightarrow]-\infty, +\infty]$ which are proper in the sense that

$$\text{dom } \varphi = \{x \in \mathcal{H} \mid \varphi(x) < +\infty\} \neq \emptyset. \tag{4}$$

Let C and D be nonempty convex subsets of \mathcal{H} . The indicator function of C is

$$\iota_C : x \mapsto \begin{cases} 0, & \text{if } x \in C; \\ +\infty, & \text{if } x \notin C. \end{cases} \tag{5}$$

If C is closed, for every $x \in \mathcal{H}$, there exists a unique point $P_C x \in C$ such that $\|x - P_C x\| = \inf_{y \in C} \|x - y\|$; $P_C x$ is called the projection of x onto C . We say that 0 lies in the strong relative interior of C , in symbol, $0 \in \text{sri } C$, if $\bigcup_{\lambda > 0} \lambda C = \overline{\text{span}} C$. In particular, if we set $C - D = \{x - y \mid (x, y) \in C \times D\}$, the inclusion $0 \in \text{sri}(C - D)$ holds in each of the following cases:

- $C - D$ is a closed vector subspace.
- $0 \in \text{int}(C - D)$.
- $C \cap \text{int } D \neq \emptyset$.
- \mathcal{H} is finite dimensional and $(\text{ri } C) \cap (\text{ri } D) \neq \emptyset$, where $\text{ri } C$ denotes the relative interior of C , i.e., its interior relative to its affine hull.

General background on convex analysis will be found in [8, 47].

3 Proximity Operators

3.1 Definition and Properties

For a detailed account of the theory of proximity operators, see [8] and the classical paper [33].

Let $\varphi \in \Gamma_0(\mathcal{H})$. For every $x \in \mathcal{H}$, the function

$$y \mapsto \varphi(y) + \frac{1}{2} \|x - y\|^2 \tag{6}$$

has a unique minimizer, which is denoted by $\text{prox}_\varphi x$ and characterized by the variational inequality

$$\begin{aligned} (\forall p \in \mathcal{H}) \quad p = \text{prox}_\varphi x \\ \Leftrightarrow (\forall y \in \mathcal{H}) \quad \langle y - p | x - p \rangle + \varphi(p) \leq \varphi(y). \end{aligned} \tag{7}$$

The proximity operator prox_φ of φ thus defined is nonexpansive, i.e.,

$$(\forall x \in \mathcal{H}) (\forall y \in \mathcal{H}) \quad \|\text{prox}_\varphi x - \text{prox}_\varphi y\| \leq \|x - y\|. \tag{8}$$

Example 1 Let C be a nonempty closed convex subset of \mathcal{H} . Then $\text{prox}_{\iota_C} = P_C$.

Other closed-form expressions for the proximity operators can be found in [3, 10, 14, 17, 18, 21, 33].

Lemma 1 [17, Proposition 11] *Let $\psi \in \Gamma_0(\mathcal{G})$, let $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$, and set $\varphi = \psi \circ L$. Suppose that $L \circ L^* = \alpha \text{Id}$, for some $\alpha \in]0, +\infty[$. Then $\varphi \in \Gamma_0(\mathcal{H})$ and*

$$\text{prox}_\varphi = \text{Id} + \frac{1}{\alpha} L^* \circ (\text{prox}_{\alpha\psi} - \text{Id}) \circ L. \tag{9}$$

3.2 Multicomponent Proximity Operators

The computation of proximity operators in the Hilbert direct sum \mathcal{H} will play a fundamental role in the next sections. Below, we provide some important situations in which this computation is explicit.

Proposition 1 *Suppose that, for every $i \in \{1, \dots, m\}$, $(e_{i,k})_{k \in \mathbb{K}}$ is an orthonormal basis of \mathcal{H}_i . Furthermore, let $(\phi_k)_{k \in \mathbb{K}}$ be functions in $\Gamma_0(\mathbb{R}^m)$ and suppose that one of the following holds.*

- (i) *For every $i \in \{1, \dots, m\}$, \mathcal{H}_i is infinite dimensional and, for every $k \in \mathbb{K}$, $\phi_k \geq \phi_k(\mathbf{0}) = 0$.*
- (ii) *For every $i \in \{1, \dots, m\}$, \mathcal{H}_i is finite dimensional.*

Set

$$f: \mathcal{H} \rightarrow]-\infty, +\infty]$$

$$x \mapsto \sum_{k \in \mathbb{K}} \phi_k(\langle x_1 | e_{1,k} \rangle, \dots, \langle x_m | e_{m,k} \rangle). \tag{10}$$

Then $f \in \Gamma_0(\mathcal{H})$ and, for every $x \in \mathcal{H}$,

$$\text{prox}_f x = \left(\sum_{k \in \mathbb{K}} \pi_{1,k} e_{1,k}, \dots, \sum_{k \in \mathbb{K}} \pi_{m,k} e_{m,k} \right), \tag{11}$$

where

$$(\forall k \in \mathbb{K}) \quad (\pi_{1,k}, \dots, \pi_{m,k}) = \text{prox}_{\phi_k}(\langle x_1 | e_{1,k} \rangle, \dots, \langle x_m | e_{m,k} \rangle). \tag{12}$$

Proof Let us set, for every $k \in \mathbb{K}$,

$$\psi_k: \mathcal{H} \rightarrow]-\infty, +\infty]$$

$$x \mapsto \phi_k(\langle x_1 | e_{1,k} \rangle, \dots, \langle x_m | e_{m,k} \rangle). \tag{13}$$

Then our assumptions imply that the functions $(\psi_k)_{k \in \mathbb{K}}$ are in $\Gamma_0(\mathcal{H})$. Under assumption (i), assuming without loss of generality that $\mathbb{K} = \mathbb{N}$, we can write $f = \sup_{K \in \mathbb{N}} \sum_{k=0}^K \psi_k$. Since lower semicontinuity and convexity are preserved under finite sums and taking suprema, it follows that f is lower semicontinuous and convex. In addition, since $f(\mathbf{0}) = 0$, we obtain $f \in \Gamma_0(\mathcal{H})$. On the other hand, under assumption 1, the sum in (10) is finite and our assumptions imply at once that $f \in \Gamma_0(\mathcal{H})$.

Now let $x \in \mathcal{H}$ and denote the Euclidean norm on \mathbb{R}^m by $|\cdot|$. Set

$$(\forall i \in \{1, \dots, m\}) (\forall k \in \mathbb{K}) \quad \xi_{i,k} = \langle x_i | e_{i,k} \rangle. \tag{14}$$

Moreover, for every $i \in \{1, \dots, m\}$, let $y_i \in \mathcal{H}_i$ and set $(\eta_{i,k})_{k \in \mathbb{K}} = ((y_i | e_{i,k}))_{k \in \mathbb{K}}$. We derive from (12) and (7)

that, for every $k \in \mathbb{K}$,

$$\sum_{i=1}^m (\eta_{i,k} - \pi_{i,k})(\xi_{i,k} - \pi_{i,k}) + \phi_k(\pi_{1,k}, \dots, \pi_{m,k}) \leq \phi_k(\eta_{1,k}, \dots, \eta_{m,k}). \tag{15}$$

Let us first assume that 1 holds. For every $k \in \mathbb{K}$ observe that, since $\mathbf{0}$ is a minimizer of ϕ_k , (7) yields $\text{prox}_{\phi_k} \mathbf{0} = \mathbf{0}$. Hence, using (12), (14), (8), and Parseval’s identity, we obtain

$$\begin{aligned} & \sum_{k \in \mathbb{K}} \sum_{i=1}^m |\pi_{i,k}|^2 \\ &= \sum_{k \in \mathbb{K}} |(\pi_{1,k}, \dots, \pi_{m,k})|^2 \\ &= \sum_{k \in \mathbb{K}} |\text{prox}_{\phi_k}(\xi_{1,k}, \dots, \xi_{m,k})|^2 \\ &= \sum_{k \in \mathbb{K}} |\text{prox}_{\phi_k}(\xi_{1,k}, \dots, \xi_{m,k}) - \text{prox}_{\phi_k} \mathbf{0}|^2 \\ &\leq \sum_{k \in \mathbb{K}} |(\xi_{1,k}, \dots, \xi_{m,k}) - \mathbf{0}|^2 \\ &= \sum_{k \in \mathbb{K}} |(\xi_{1,k}, \dots, \xi_{m,k})|^2 \\ &= \sum_{k \in \mathbb{K}} \sum_{i=1}^m |\xi_{i,k}|^2 \\ &= \sum_{i=1}^m \|x_i\|^2. \end{aligned} \tag{16}$$

Therefore, $(\forall i \in \{1, \dots, m\}) \sum_{k \in \mathbb{K}} |\pi_{i,k}|^2 < +\infty$. Consequently, we can define

$$(\forall i \in \{1, \dots, m\}) \quad z_i = \sum_{k \in \mathbb{K}} \pi_{i,k} e_{i,k}, \tag{17}$$

which is equally well defined under assumption (ii) as \mathbb{K} is then finite. It remains to show that $(z_i)_{1 \leq i \leq m} = \text{prox}_f(x_1, \dots, x_m)$. Summing over k in (15) yields

$$\sum_{k \in \mathbb{K}} \sum_{i=1}^m (\eta_{i,k} - \pi_{i,k})(\xi_{i,k} - \pi_{i,k}) + \sum_{k \in \mathbb{K}} \phi_k(\pi_{1,k}, \dots, \pi_{m,k}) \leq \sum_{k \in \mathbb{K}} \phi_k(\eta_{1,k}, \dots, \eta_{m,k}) \tag{18}$$

and, therefore,

$$\sum_{i=1}^m (y_i - z_i | x_i - z_i) + g(z_1, \dots, z_m) \leq g(y_1, \dots, y_m). \tag{19}$$

In view of (7), the proof is complete. \square

Proposition 2 For every $j \in \{1, \dots, q\}$, let \mathcal{G}_j be a real Hilbert space, let $\varphi_j \in \Gamma_0(\mathcal{G}_j)$, and, for every $i \in \{1, \dots, m\}$, let $L_{j,i} \in \mathcal{B}(\mathcal{H}_i, \mathcal{G}_j)$. Set

$$f : \mathcal{H} \rightarrow]-\infty, +\infty]$$

$$x \mapsto \sum_{j=1}^q \varphi_j \left(\sum_{i=1}^m L_{j,i} x_i \right) \tag{20}$$

and suppose that, for every $j \in \{1, \dots, q\}$, there exists $\alpha_j \in]0, +\infty[$ such that

$$(\forall k \in \{1, \dots, q\}) \sum_{i=1}^m L_{j,i} \circ L_{k,i}^* = \begin{cases} \alpha_j \text{Id}, & \text{if } j = k; \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$

Then $f \in \Gamma_0(\mathcal{H})$ and, for every $x \in \mathcal{H}$,

$$\text{prox}_f x = (p_1, \dots, p_m) \tag{22}$$

where, for every $i \in \{1, \dots, m\}$,

$$p_i = x_i + \sum_{j=1}^q \alpha_j^{-1} L_{j,i}^* \text{prox}_{\alpha_j \varphi_j} \left(\sum_{k=1}^m L_{j,k} x_k \right) - \sum_{j=1}^q \alpha_j^{-1} L_{j,i}^* \sum_{k=1}^m L_{j,k} x_k. \tag{23}$$

Proof Let us denote by \mathcal{G} the product space $\mathcal{G}_1 \times \dots \times \mathcal{G}_q$ equipped with the usual vector space structure and the scalar product

$$(y, z) \mapsto \sum_{j=1}^q \alpha_j^{-1} \langle y_j \mid z_j \rangle. \tag{24}$$

We can write $f = g \circ L$, where

$$g : \mathcal{G} \rightarrow]-\infty, +\infty] \tag{25}$$

$$y \mapsto \sum_{j=1}^q \varphi_j(y_j)$$

and $L \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ is defined by

$$L : \mathcal{H} \rightarrow \mathcal{G} \tag{26}$$

$$x \mapsto \left(\sum_{i=1}^m L_{1,i} x_i, \dots, \sum_{i=1}^m L_{q,i} x_i \right).$$

It follows from (24) that, for every $(x, y) \in \mathcal{H} \times \mathcal{G}$,

$$\langle Lx \mid y \rangle = \sum_{j=1}^q \alpha_j^{-1} \left\langle \sum_{i=1}^m L_{j,i} x_i \mid y_j \right\rangle$$

$$= \sum_{j=1}^q \alpha_j^{-1} \sum_{i=1}^m \langle x_i \mid L_{j,i}^* y_j \rangle = \sum_{i=1}^m \left\langle x_i \mid \sum_{j=1}^q \alpha_j^{-1} L_{j,i}^* y_j \right\rangle, \tag{27}$$

from which we deduce that the adjoint of L is

$$L^* : \mathcal{G} \rightarrow \mathcal{H} \tag{28}$$

$$y \mapsto \left(\sum_{j=1}^q \alpha_j^{-1} L_{j,1}^* y_j, \dots, \sum_{j=1}^q \alpha_j^{-1} L_{j,m}^* y_j \right).$$

We then get from (21) that $L \circ L^* = \text{Id}$. Hence, Lemma 1 implies that $f = g \circ L \in \Gamma_0(\mathcal{H})$ and that

$$\text{prox}_{g \circ L} = \text{Id} + L^* \circ (\text{prox}_g - \text{Id}) \circ L. \tag{29}$$

In addition, it follows from (25) and (24) that, for every $y \in \mathcal{G}$,

$$\text{prox}_g y = (\text{prox}_{\alpha_1 \varphi_1} y_1, \dots, \text{prox}_{\alpha_q \varphi_q} y_q). \tag{30}$$

Altogether, (26), (28), (29), and (30) yield (22)–(23). \square

Corollary 1 Let $\varphi \in \Gamma_0(\mathcal{G})$ and, for every $i \in \{1, \dots, m\}$, let $L_i \in \mathcal{B}(\mathcal{H}_i, \mathcal{G})$. Set

$$f : \mathcal{H} \rightarrow]-\infty, +\infty] \tag{31}$$

$$x \mapsto \varphi \left(\sum_{i=1}^m L_i x_i \right)$$

and suppose that there exists $\alpha \in]0, +\infty[$ such that

$$\sum_{i=1}^m L_i \circ L_i^* = \alpha \text{Id}. \tag{32}$$

Then $f \in \Gamma_0(\mathcal{H})$ and, for every $x \in \mathcal{H}$, $\text{prox}_f x = (p_1, \dots, p_m)$ where, for every $i \in \{1, \dots, m\}$,

$$p_i = x_i + \alpha^{-1} L_i^* \text{prox}_{\alpha \varphi} \left(\sum_{k=1}^m L_k x_k \right) - \alpha^{-1} L_i^* \sum_{k=1}^m L_k x_k. \tag{33}$$

Proof Set $q = 1$ in Proposition 2. \square

Proposition 3 Suppose that \mathcal{G} has finite dimension K , let $(\phi_k)_{1 \leq k \leq K}$ be functions in $\Gamma_0(\mathbb{R})$, and let $(e_k)_{1 \leq k \leq K}$ be an orthonormal basis of \mathcal{G} . For every $i \in \{1, \dots, m\}$, let $L_i \in \mathcal{B}(\mathcal{H}_i, \mathcal{G})$ and suppose that there exists $\{\alpha_k\}_{1 \leq k \leq K} \subset \mathbb{R}$

$]0, +\infty[$ such that

$$(\forall y \in \mathcal{G}) \quad \sum_{i=1}^m L_i L_i^* y = \sum_{k=1}^K \alpha_k \langle y | e_k \rangle e_k. \tag{34}$$

Set

$$f : \mathcal{H} \rightarrow]-\infty, +\infty] : x \mapsto \sum_{k=1}^K \phi_k \left(\left\langle \sum_{j=1}^m L_j x_j \mid e_k \right\rangle \right) \tag{35}$$

and, for every $k \in \{1, \dots, K\}$,

$$\pi_k = \frac{1}{\alpha_k} \operatorname{prox}_{\alpha_k \phi_k} \left(\left\langle \sum_{j=1}^m L_j x_j \mid e_k \right\rangle \right). \tag{36}$$

Then $f \in \Gamma_0(\mathcal{H})$ and, for every $x \in \mathcal{H}$, $\operatorname{prox}_f x = (p_i)_{1 \leq i \leq m}$ where, for every $i \in \{1, \dots, m\}$,

$$p_i = x_i + L_i^* \sum_{k=1}^K \left(\pi_k - \frac{1}{\alpha_k} \sum_{j=1}^m \langle L_j x_j | e_k \rangle \right) e_k. \tag{37}$$

Proof For every $j \in \{1, \dots, K\}$, set $\mathcal{G}_j = \mathbb{R}$, $\varphi_j = \phi_j$, and

$$(\forall i \in \{1, \dots, m\}) \quad L_{j,i} : \mathcal{H}_i \rightarrow \mathcal{G}_j : x \mapsto \langle L_i x | e_j \rangle, \tag{38}$$

hence

$$L_{j,i}^* : \mathcal{G}_j \rightarrow \mathcal{H}_i : \xi \mapsto \xi L_i^* e_j. \tag{39}$$

Thus, for every j and k in $\{1, \dots, K\}$ and every $\xi \in \mathbb{R}$, we derive from (34) that

$$\begin{aligned} \sum_{i=1}^m L_{j,i} L_{k,i}^* \xi &= \sum_{i=1}^m L_{j,i} \xi L_i^* e_k \\ &= \xi \sum_{i=1}^m \langle (L_i L_i^*) e_k | e_j \rangle \\ &= \xi \left\langle \sum_{l=1}^K \alpha_l \langle e_k | e_l \rangle e_l \mid e_j \right\rangle \\ &= \xi \sum_{l=1}^K \alpha_l \langle e_k | e_l \rangle \langle e_l | e_j \rangle. \end{aligned} \tag{40}$$

Therefore, for every $j \in \{1, \dots, K\}$, (21) is satisfied. In turn, Proposition 2 with $q = K$ guarantees that $f \in \Gamma_0(\mathcal{H})$, and (23) reduces to (37). \square

4 Multicomponent Proximal Algorithms

We present several algorithms for solving Problem 1 under various assumptions on the functions involved. Most of

these algorithms are tolerant to errors in the computation of proximal points and gradients. To quantify the amount of error which is tolerated, it will be convenient to use the following notation: given two sequences $(x_n)_{n \in \mathbb{N}}$ and $(y_n)_{n \in \mathbb{N}}$ in \mathcal{H} ,

$$\left[(\forall n \in \mathbb{N}) \quad x_n \approx y_n \right] \Leftrightarrow \sum_{n \in \mathbb{N}} \|x_n - y_n\| < +\infty. \tag{41}$$

4.1 Forward-Backward Splitting

Problem 2 In Problem 1, suppose that $p = 2$ and that f_2 is differentiable on \mathcal{H} with a β -Lipschitz continuous gradient for some $\beta \in]0, +\infty[$. Hence, the problem is to

$$\underset{x_1 \in \mathcal{H}_1, \dots, x_m \in \mathcal{H}_m}{\text{minimize}} \quad f_1(x_1, \dots, x_m) + f_2(x_1, \dots, x_m), \tag{42}$$

under the assumption that solutions exist.

The particular case when f_1 is a separable sum and f_2 involves a linear mixture of the variables was investigated in [10]. The following result addresses the general case; it implicitly assumes that the proximity operator of f_1 can be computed to within a quantifiable error.

Theorem 1 Let $(x_{1,n})_{n \in \mathbb{N}}, \dots, (x_{m,n})_{n \in \mathbb{N}}$ be sequences generated by the following routine.

Initialization

$$\left[\begin{array}{l} \varepsilon \in]0, \min\{1, 1/\beta\}[\\ \text{For } i = 1, \dots, m \\ \lfloor x_{i,0} \in \mathcal{H}_i \end{array} \right.$$

For $n = 0, 1, \dots$

$$\left[\begin{array}{l} (y_{i,n})_{1 \leq i \leq m} \approx \nabla f_2(x_{i,n})_{1 \leq i \leq m} \\ \gamma_n \in [\varepsilon, (2/\beta) - \varepsilon] \\ (u_{i,n})_{1 \leq i \leq m} \approx \operatorname{prox}_{\gamma_n f_1}(x_{i,n} - \gamma_n y_{i,n})_{1 \leq i \leq m} \\ \lambda_n \in [\varepsilon, 1] \\ \text{For } i = 1, \dots, m \\ \lfloor x_{i,n+1} = x_{i,n} + \lambda_n (u_{i,n} - x_{i,n}) \end{array} \right.$$

Then, for every $i \in \{1, \dots, m\}$, $(x_{i,n})_{n \in \mathbb{N}}$ converges weakly to a point $x_i \in \mathcal{H}_i$. Moreover, $(x_i)_{1 \leq i \leq m}$ is a solution to Problem 2.

Proof Apply [21, Theorem 3.4(i)] in \mathcal{H} and use (3). \square

Remark 1

- (i) Multicomponent version of variants of the above forward-backward algorithm such as the Nesterov-like first-order methods [9, 34, 45] can be obtained by similar reformulations in \mathcal{H} . However, for these methods,

convergence of the iterates $((x_{i,n})_{1 \leq i \leq m})_{n \in \mathbb{N}}$ to a solution to Problem 2 is not guaranteed, even in a finite-dimensional setting.

- (ii) Strong convergence conditions in Theorem 1 can be derived from [21, Theorem 3.4(iv)].

4.2 Douglas-Rachford Splitting

In this section, we relax the assumption of smoothness on f_2 and assume that its proximity operator is implementable to within a quantifiable error.

Problem 3 In Problem 1, suppose that $p = 2$ and that

$$0 \in \text{sri}(\text{dom } f_1 - \text{dom } f_2). \tag{43}$$

Hence, the problem is to

$$\underset{x_1 \in \mathcal{H}_1, \dots, x_m \in \mathcal{H}_m}{\text{minimize}} \quad f_1(x_1, \dots, x_m) + f_2(x_1, \dots, x_m), \tag{44}$$

under the assumption that solutions exist.

Theorem 2 Let $(x_{1,n})_{n \in \mathbb{N}}, \dots, (x_{m,n})_{n \in \mathbb{N}}$ be sequences generated by the following routine.

Initialization

$$\left[\begin{array}{l} \varepsilon \in]0, 1[\\ \gamma \in]0, +\infty[\\ \text{For } i = 1, \dots, m \\ \lfloor x_{i,0} \in \mathcal{H}_i \end{array} \right.$$

For $n = 0, 1, \dots$

$$\left[\begin{array}{l} (y_{i,n})_{1 \leq i \leq m} \approx \text{prox}_{\gamma f_2}(x_{i,n})_{1 \leq i \leq m} \\ (u_{i,n})_{1 \leq i \leq m} \approx \text{prox}_{\gamma f_1}(2y_{i,n} - x_{i,n})_{1 \leq i \leq m} \\ \lambda_n \in [\varepsilon, 2 - \varepsilon] \\ \text{For } i = 1, \dots, m \\ \lfloor x_{i,n+1} = x_{i,n} + \lambda_n(u_{i,n} - y_{i,n}). \end{array} \right.$$

Then, for every $i \in \{1, \dots, m\}$, $(x_{i,n})_{n \in \mathbb{N}}$ converges weakly to a point $x_i \in \mathcal{H}_i$. Moreover, $\text{prox}_{\gamma f_2}(x_1, \dots, x_m)$ is a solution to Problem 3 and $((y_{i,n})_{1 \leq i \leq m})_{n \in \mathbb{N}}$ converges weakly to $\text{prox}_{\gamma f_2}(x_1, \dots, x_m)$.

Proof For the first two claims, apply [17, Theorem 20] in \mathcal{H} and use (3). For the weak convergence claim, combine the results of [8, Corollary 27.4] and [39]. \square

Remark 2

- (i) Strong convergence conditions in Theorem 2 can be derived from [16, Theorem 2.1(ii)].
- (ii) If \mathcal{H} is finite dimensional, the qualification condition (43) reduces to

$$(\text{ri dom } f_1) \cap (\text{ri dom } f_2) \neq \emptyset. \tag{45}$$

4.3 Parallel Proximal Algorithm (PPXA)

The algorithm presented in this section aims at solving Problem 1 under minimal technical assumptions. Its cost of implementation depends on the ease of (approximate) computation of the individual proximity operators.

Problem 4 In Problem 1, suppose that

$$0 \in \text{sri}(\mathbf{D} - \text{dom } f_1 \times \dots \times \text{dom } f_p) \tag{46}$$

where $\mathbf{D} = \{(x, \dots, x) \mid x \in \mathcal{H}\}$. The problem is to

$$\underset{x_1 \in \mathcal{H}_1, \dots, x_m \in \mathcal{H}_m}{\text{minimize}} \quad \sum_{k=1}^p f_k(x_1, \dots, x_m), \tag{47}$$

under the assumption that solutions exist.

In [1], a particular instance of Problem 4 in finite dimensional spaces is considered. It is approached via the alternating direction method of multipliers. The algorithm used below is an application of the PPXA algorithm proposed in [19] that allows us to address the general case.

Theorem 3 Let $(x_{1,n})_{n \in \mathbb{N}}, \dots, (x_{m,n})_{n \in \mathbb{N}}$ be sequences generated by the following routine.

Initialization

$$\left[\begin{array}{l} \varepsilon \in]0, 1[\\ \gamma \in]0, +\infty[\\ \{\omega_k\}_{1 \leq k \leq p} \subset]0, 1] \quad \text{and} \quad \sum_{k=1}^p \omega_k = 1 \\ \text{For } i = 1, \dots, m \\ \left[\begin{array}{l} \text{For } k = 1, \dots, p \\ \lfloor y_{i,k,0} \in \mathcal{H}_i \\ \lfloor x_{i,0} = \sum_{k=1}^p \omega_k y_{i,k,0} \end{array} \right. \end{array} \right.$$

For $n = 0, 1, \dots$

$$\left[\begin{array}{l} \text{For } k = 1, \dots, p \\ \lfloor (u_{i,k,n})_{1 \leq i \leq m} \approx \text{prox}_{\gamma f_k / \omega_k}(y_{i,k,n})_{1 \leq i \leq m} \\ \text{For } i = 1, \dots, m \\ \left[\begin{array}{l} s_{i,n} = \sum_{k=1}^p \omega_k u_{i,k,n} \\ \lambda_n \in [\varepsilon, 2 - \varepsilon] \\ \text{For } k = 1, \dots, p \\ \lfloor y_{i,k,n+1} = y_{i,k,n} + \lambda_n(2s_{i,n} - x_{i,n} - u_{i,k,n}) \\ \lfloor x_{i,n+1} = x_{i,n} + \lambda_n(s_{i,n} - x_{i,n}) \end{array} \right. \end{array} \right.$$

Then, for every $i \in \{1, \dots, m\}$, $(x_{i,n})_{n \in \mathbb{N}}$ converges weakly to a point $x_i \in \mathcal{H}_i$. Moreover, (x_1, \dots, x_m) is a solution to Problem 4.

Proof Apply [19, Theorem 3.4] in \mathcal{H} and use (3). \square

Remark 3 Suppose that \mathcal{H} is finite dimensional and that

$$\bigcap_{k=1}^p \text{ri dom } f_k \neq \emptyset. \tag{48}$$

Then it follows from [19, Proposition 3.6(vi)] that the qualification condition (46) is satisfied.

4.4 Dykstra-like Splitting

We consider instances of Problem 1 in which f_p is a simple quadratic function.

Problem 5 In Problem 1, suppose that $p \geq 3$, that

$$\bigcap_{k=1}^{p-1} \text{dom } f_k \neq \emptyset, \tag{49}$$

and that $f_p: \mathbf{x} \mapsto (p-1) \sum_{i=1}^m \|x_i - z_i\|^2/2$, where $\mathbf{z} \in \mathcal{H}$. Hence, the problem is to

$$\begin{aligned} \text{minimize}_{x_1 \in \mathcal{H}_1, \dots, x_m \in \mathcal{H}_m} & \sum_{k=1}^{p-1} f_k(x_1, \dots, x_m) \\ & + \frac{p-1}{2} \sum_{i=1}^m \|x_i - z_i\|^2. \end{aligned} \tag{50}$$

Set $f = \sum_{k=1}^{p-1} f_k / (p-1)$. Then it follows from (49) that $f \in \Gamma_0(\mathcal{H})$. Hence, in view of (3), Problem 5 admits a unique solution, namely $\text{prox}_f \mathbf{z}$.

Theorem 4 Let $(x_{1,n})_{n \in \mathbb{N}}, \dots, (x_{m,n})_{n \in \mathbb{N}}$ be sequences generated by the following routine.

Initialization

$$\left[\begin{array}{l} \text{For } i = 1, \dots, m \\ \quad x_{i,0} = z_i \\ \quad \text{For } k = 1, \dots, p-1 \\ \quad \quad y_{i,k,0} = x_{i,0} \end{array} \right.$$

For $n = 0, 1, \dots$

$$\left[\begin{array}{l} \text{For } k = 1, \dots, p-1 \\ \quad \lfloor (u_{i,k,n})_{1 \leq i \leq m} = \text{prox}_{f_k} (y_{i,k,n})_{1 \leq i \leq m} \\ \text{For } i = 1, \dots, m \\ \quad x_{i,n+1} = \frac{1}{p-1} \sum_{k=1}^{p-1} u_{i,k,n} \\ \quad \text{For } k = 1, \dots, p-1 \\ \quad \quad \lfloor y_{i,k,n+1} = x_{i,n+1} + y_{i,k,n} - u_{i,k,n}. \end{array} \right. \tag{51}$$

Then, for every $i \in \{1, \dots, m\}$, $(x_{i,n})_{n \in \mathbb{N}}$ converges strongly to a point $x_i \in \mathcal{H}_i$. Moreover, (x_1, \dots, x_m) is a solution to Problem 5.

Proof Apply [16, Theorem 4.2] in \mathcal{H} and use (3). \square

Remark 4 Suppose that (49) is replaced by the stronger condition (46) (applied to the functions $(f_k)_{1 \leq k \leq p-1}$). Then it follows from [16, Theorem 3.3] that the conclusion of the above theorem remains valid if the proximity operators are implemented approximately in (51).

5 Applications to Image Decomposition and Recovery

In this section, we apply the algorithms proposed in Sect. 4 to stereoscopic image restoration, multispectral imaging, and image decomposition.

5.1 Stereoscopic Image Restoration

5.1.1 Problem Formulation

We consider the problem of restoring a pair of N -pixel stereoscopic images $\bar{x}_1 \in \mathbb{R}^N$ and $\bar{x}_2 \in \mathbb{R}^N$, which correspond to the left and the right views of the same scene. For a given value of the disparity field, the disparity compensation process between the two images is modeled as

$$\bar{x}_1 = D\bar{x}_2 + v, \tag{52}$$

where the matrix D is in $\mathbb{R}^{N \times N}$ [38] and where v stands for modeling errors. The observations consist of degraded versions

$$z_1 = L_1 \bar{x}_1 + w_1 \quad \text{and} \quad z_2 = L_2 \bar{x}_2 + w_2 \tag{53}$$

of \bar{x}_1 and \bar{x}_2 , respectively, where the matrices $L_1 \in \mathbb{R}^{N \times N}$ and $L_2 \in \mathbb{R}^{N \times N}$ model the data acquisition process, and where w_1 and w_2 are mutually independent Gaussian noise vectors with independent components which are $\mathcal{N}(0, \sigma_1^2)$ - and $\mathcal{N}(0, \sigma_2^2)$ -distributed, respectively. In addition, it is assumed that the decompositions of \bar{x}_1 and \bar{x}_2 in orthonormal bases $(e_{1,k})_{1 \leq k \leq N}$ and $(e_{2,k})_{1 \leq k \leq N}$, respectively, of \mathbb{R}^N are sparse. For every $k \in \{1, \dots, N\}$, functions $\phi_{1,k} \in \Gamma_0(\mathbb{R})$ and $\phi_{2,k} \in \Gamma_0(\mathbb{R})$ are used to promote the sparsity of the decompositions [18, 22]. The following variational formulation is consistent with the above hypotheses and models.

Problem 6 Let $\vartheta \in [0, +\infty[$. The objective is to

$$\begin{aligned} \text{minimize}_{x_1 \in \mathbb{R}^N, x_2 \in \mathbb{R}^N} & \sum_{k=1}^N \phi_{1,k}(\langle x_1 | e_{1,k} \rangle) \\ & + \sum_{k=1}^N \phi_{2,k}(\langle x_2 | e_{2,k} \rangle) \\ & + \frac{1}{2\sigma_1^2} \|L_1 x_1 - z_1\|^2 \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{2\sigma_2^2} \|L_2x_2 - z_2\|^2 \\
 & + \frac{\vartheta}{2} \|x_1 - Dx_2\|^2. \tag{54}
 \end{aligned}$$

We can formulate Problem 6 as an instance of Problem 1 with $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{R}^N$ and $m = 2$ functions, namely

$$\begin{aligned}
 f_1: (x_1, x_2) \mapsto & \sum_{k=1}^N \phi_{1,k}(\langle x_1 | e_{1,k} \rangle) \\
 & + \sum_{k=1}^N \phi_{2,k}(\langle x_2 | e_{2,k} \rangle) \tag{55}
 \end{aligned}$$

and

$$\begin{aligned}
 f_2: (x_1, x_2) \mapsto & \frac{1}{2\sigma_1^2} \|L_1x_1 - z_1\|^2 + \frac{1}{2\sigma_2^2} \|L_2x_2 - z_2\|^2 \\
 & + \frac{\vartheta}{2} \|x_1 - Dx_2\|^2. \tag{56}
 \end{aligned}$$

Proposition 4 *Let x_1 and x_2 be arbitrary vectors in \mathbb{R}^N . Then f_2 is differentiable at (x_1, x_2) and*

$$\begin{aligned}
 \nabla f_2(x_1, x_2) = & \left(\frac{1}{\sigma_1^2} L_1^\top (L_1x_1 - z_1) + \vartheta(x_1 - Dx_2), \right. \\
 & \left. \frac{1}{\sigma_2^2} L_2^\top (L_2x_2 - z_2) + \vartheta D^\top (Dx_2 - x_1) \right). \tag{57}
 \end{aligned}$$

Moreover, ∇f_2 is β -Lipschitz continuous, with

$$\beta = \max\{\sigma_1^{-2} \|L_1\|^2, \sigma_2^{-2} \|L_2\|^2\} + \vartheta(1 + \|D\|^2). \tag{58}$$

Proof The expression (57) follows from straightforward calculus. Now set

$$L = \begin{bmatrix} \sigma_1^{-1} L_1 & [0] \\ [0] & \sigma_2^{-1} L_2 \end{bmatrix} \quad \text{and} \quad M = \sqrt{\vartheta} \begin{bmatrix} I & -D \\ [0] & [0] \end{bmatrix}. \tag{59}$$

Then, using matrix notation, we can write

$$\nabla f_2 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (L^\top L + M^\top M) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - L^\top \begin{bmatrix} \sigma_1^{-1} z_1 \\ \sigma_2^{-1} z_2 \end{bmatrix}. \tag{60}$$

Hence, a Lipschitz constant of ∇f_2 is $\|L^\top L + M^\top M\|$, where $\|\cdot\|$ denotes the spectral norm. To obtain a tractable bound, we observe that

$$\begin{aligned}
 \|L^\top L + M^\top M\| & \leq \|L^\top L\| + \|M^\top M\| \\
 & = \|L\|^2 + \|M\|^2
 \end{aligned}$$

$$= \|L\|^2 + \vartheta(1 + \|D\|^2). \tag{61}$$

Now set $\mathbf{x} = [x_1 \ x_2]^\top$. Then

$$\begin{aligned}
 \|\mathbf{L}\mathbf{x}\|^2 & = \sigma_1^{-2} \|L_1x_1\|^2 + \sigma_2^{-2} \|L_2x_2\|^2 \\
 & \leq \sigma_1^{-2} \|L_1\|^2 \|x_1\|^2 + \sigma_2^{-2} \|L_2\|^2 \|x_2\|^2 \\
 & \leq \max\{\sigma_1^{-2} \|L_1\|^2, \sigma_2^{-2} \|L_2\|^2\} \|\mathbf{x}\|^2. \tag{62}
 \end{aligned}$$

Hence, $\|\mathbf{L}\|^2 \leq \max\{\sigma_1^{-2} \|L_1\|^2, \sigma_2^{-2} \|L_2\|^2\}$ and (61) yields $\|\mathbf{L}^\top \mathbf{L} + \mathbf{M}^\top \mathbf{M}\| \leq \beta$. \square

In view of Proposition 4, Problem 6 can be solved by the forward-backward algorithm (see Theorem 1).

5.1.2 Numerical Experiments

Experimental results are displayed in Fig. 1 for stereoscopic images of size 256×256 ($N = 256^2$). In this example, L_1 and L_2 are periodic convolution operators with motion kernel blur of sizes 7×7 and 3×3 , respectively. This kind of blur was considered in a related context in [36]. A white Gaussian noise is added corresponding to a blurred signal-to-noise-ratio (BSNR) of 21.6 dB for z_1 and 21.8 dB for z_2 (the BSNR is defined as $10 \log_{10}(\|L_i \bar{x}_i\|^2 / (N \sigma_i^2))$). In addition, $(e_{1,k})_{1 \leq k \leq N}$ and $(e_{2,k})_{1 \leq k \leq N}$ are symmetlic wavelet orthonormal bases (length 6) over 2 resolution levels. For every $k \in \{1, \dots, N\}$, $\phi_{1,k} = \mu_{1,k} \cdot |^{p_{1,k}}$ and $\phi_{2,k} = \mu_{2,k} \cdot |^{p_{2,k}}$, where $\{\mu_{1,k}, \mu_{2,k}\} \subset]0, +\infty[$ and $\{p_{1,k}, p_{2,k}\} \subset [1, +\infty[$.

The operators $(\text{prox}_{\phi_{1,k}})_{1 \leq k \leq N}$ and $(\text{prox}_{\phi_{2,k}})_{1 \leq k \leq N}$ can be calculated explicitly [14, Examples 4.2 and 4.4]. The proximity operator of f_1 can thus be deduced from Proposition 1, the separability of this function, and [21, Lemma 2.8 and 2.9]. For every $k \in \{1, \dots, N\}$, the values of $\mu_{1,k}$, $\mu_{2,k}$, $p_{1,k}$, and $p_{2,k}$ are chosen using a maximum likelihood approach in a subband-adaptive manner with $p_{1,k}$ and $p_{2,k}$ in $\{1, 4/3, 3/2, 2\}$. The value of ϑ is selected so as to maximize the signal-to-noise-ratio (SNR). The SNR between an image y and the original image \bar{y} is defined as $20 \log_{10}(\|\bar{y}\| / \|y - \bar{y}\|)$. In our experiments we also propose to compare the restored images in terms of structural similarity (SSIM) [43]. The SSIM takes on values from -1 to 1 . The value 1 is achieved for two identical images. The disparity map has been estimated by using the method described in [32]. Note that the existence of a solution to Problem 6 is secured by the fact that $f_1 + f_2$ is a coercive function in $\Gamma_0(\mathbb{R}^N \oplus \mathbb{R}^N)$ [21, Propositions 3.1(i) and 5.15(i)]. Thus, Problem 6 is a special case of Problem 2. In this context, setting $\lambda_n \equiv 1$, the forward-backward algorithm assumes the following form.

Fig. 1 Stereoscopic image restoration

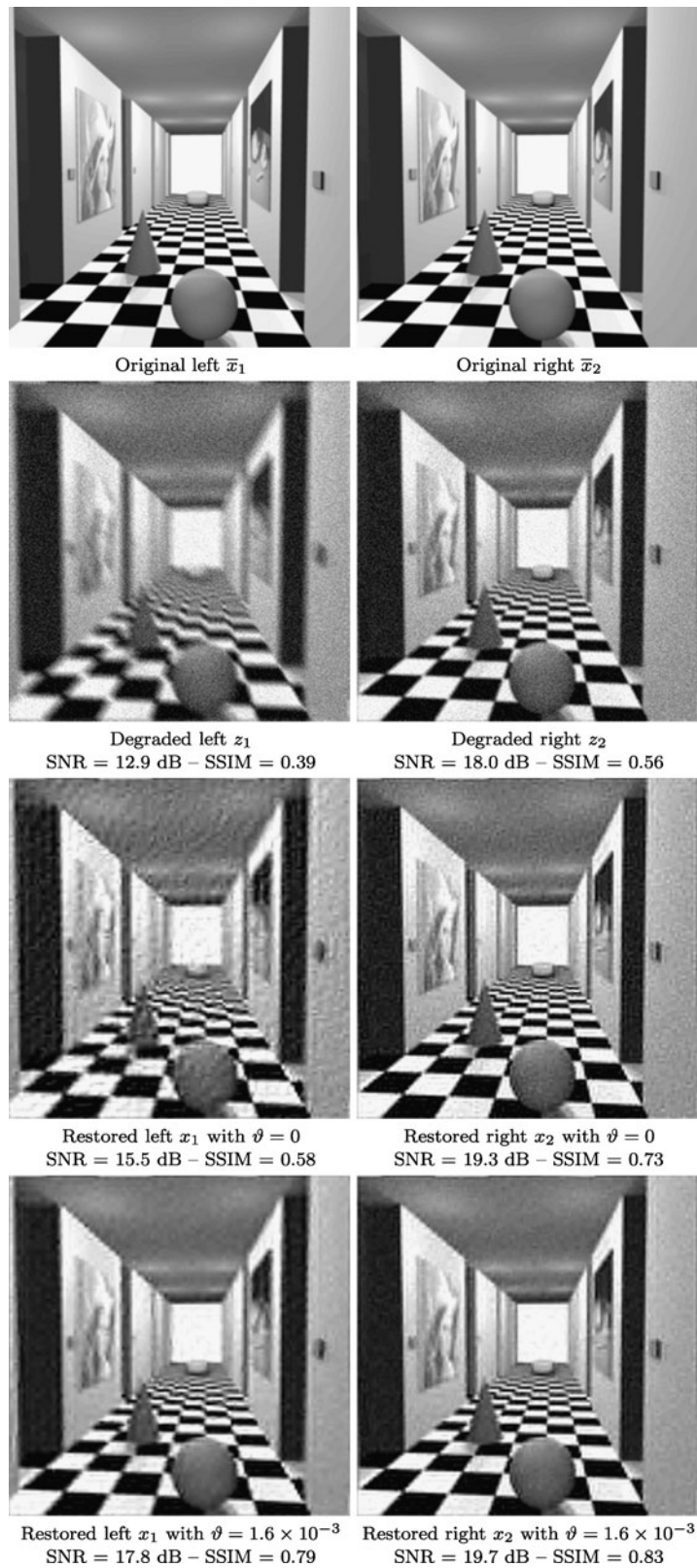
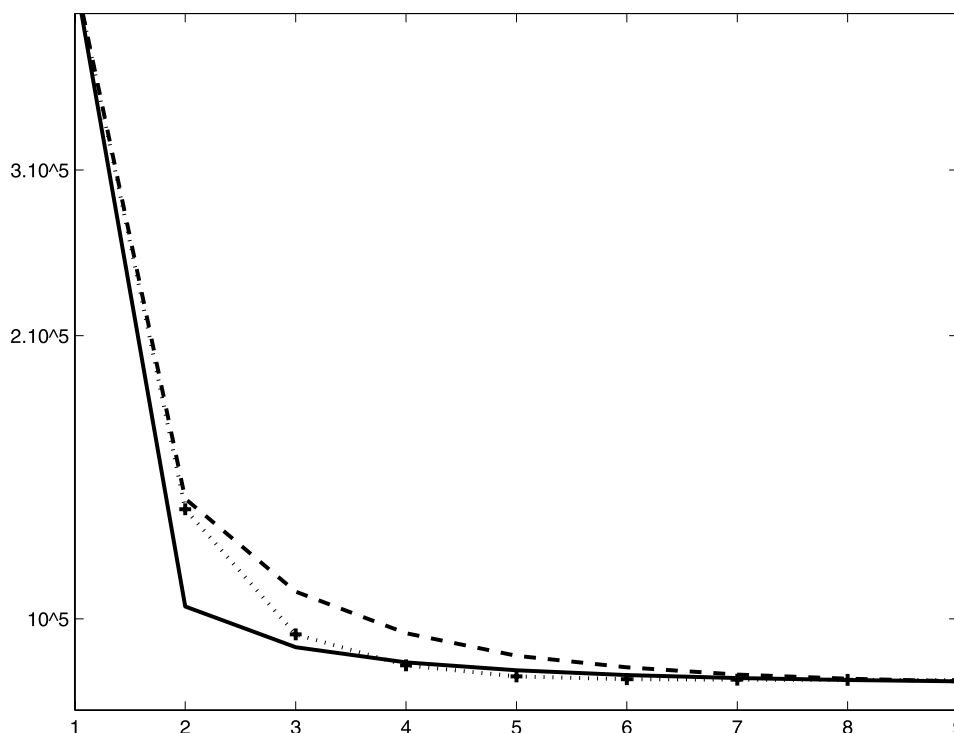


Fig. 2 Convergence of the objective function in Problem 6 for the forward-backward algorithm (solid line), the Nesterov algorithm (dotted line), and the Beck-Teboulle algorithm (dashed line) versus iteration number



Initialization

$$\begin{cases} \sigma_1 = \sigma_2 = 12 \\ \vartheta = 0 \quad \text{or} \quad \vartheta = 1.6 \times 10^{-3} \\ \gamma = 1.9 / (\max\{\sigma_1^{-2} \|L_1\|^2, \sigma_2^{-2} \|L_2\|^2\} + \vartheta(1 + \|D\|^2)) \\ x_{1,0} = z_1 \\ x_{2,0} = z_2 \end{cases}$$

For $n = 0, 1, \dots$

$$\begin{cases} y_{1,n} = \sigma_1^{-2} L_1^\top (L_1 x_{1,n} - z_1) + \vartheta (x_{1,n} - D x_{2,n}) \\ y_{2,n} = \sigma_2^{-2} L_2^\top (L_2 x_{2,n} - z_2) - \vartheta D^\top (x_{1,n} - D x_{2,n}) \\ x_{1,n+1} = \sum_{k=1}^N (\text{prox}_{\gamma \phi_{1,k}}(x_{1,n} - \gamma y_{1,n} \mid e_{1,k})) e_{1,k} \\ x_{2,n+1} = \sum_{k=1}^N (\text{prox}_{\gamma \phi_{2,k}}(x_{2,n} - \gamma y_{2,n} \mid e_{2,k})) e_{2,k} \end{cases}$$

When $\vartheta = 0$, there is no coupling between the left and right views (images in the third row of Fig. 1). As can be observed in Fig. 1, the coupling term leads to a significant improvement of the restoration, especially for the most degraded image (bottom-left image).

Using $\vartheta = 1.6 \times 10^{-3}$, we compare the forward-backward algorithm of Theorem 1 (implemented with $\lambda_n \equiv 1$ and $\gamma_n \equiv 1.99/\beta$) to a multicomponent version of the Beck-Teboulle algorithm [9] and a multicomponent version of the Nesterov algorithm [35]. Although, contrary to the forward-backward algorithm, the Beck-Teboulle and Nesterov algorithms do not insure convergence of the iterates, they are known to provide a theoretically optimal convergence rate

for the objective function. However, in this example, their performance appear to be quite comparable on that score (see Fig. 2).

5.2 Multispectral Image Denoising

5.2.1 Problem Formulation

A common multispectral image processing problem is to denoise m images $(\bar{y}_i)_{1 \leq i \leq m}$ in \mathbb{R}^N from noisy observations $(z_i)_{1 \leq i \leq m}$ given by

$$(\forall i \in \{1, \dots, m\}) \quad z_i = \bar{y}_i + w_i, \tag{63}$$

where $(w_i)_{1 \leq i \leq m}$ are realizations of mutually independent zero-mean white Gaussian noise processes with respective variances $(\sigma_i^2)_{1 \leq i \leq m}$. Early methods for multispectral image recovery are described in [28]. A tutorial on wavelet-based multispectral denoising can be found in [15].

To solve this denoising problem, we assume that, for every $i \in \{1, \dots, m\}$, \bar{y}_i satisfies some constraint represented by a nonempty closed convex set $C_i \subset \mathbb{R}^N$, and that it admits a sparse decomposition in an orthonormal basis $(e_{i,k})_{1 \leq k \leq N}$ of \mathbb{R}^N . In addition, similarities between the images are promoted by penalizing a distance between their components in some orthonormal basis $(b_k)_{1 \leq k \leq N}$ of \mathbb{R}^N . These considerations lead to the variational problem

$$\underset{y_1 \in C_1, \dots, y_m \in C_m}{\text{minimize}} \quad \sum_{i=1}^m \frac{1}{2\sigma_i^2} \|y_i - z_i\|^2$$

$$\begin{aligned}
 & + \sum_{i=1}^m \sum_{k=1}^N \tilde{\mu}_{i,k} |\langle y_i \mid e_{i,k} \rangle| \\
 & + \sum_{i=1}^{m-1} \sum_{j=i+1}^m \tilde{\vartheta}_{i,j} \sum_{k=1}^N |\langle y_i - y_j \mid b_k \rangle|
 \end{aligned} \tag{64}$$

where, for every $i \in \{1, \dots, m\}$, $\{\tilde{\mu}_{i,k}\}_{1 \leq k \leq N} \subset]0, +\infty[$ and $\{\tilde{\vartheta}_{i,j}\}_{i+1 \leq j \leq m} \subset]0, +\infty[$. After appropriate rescaling of the variables, this problem can be reformulated as follows.

Problem 7 For every $i \in \{1, \dots, m\}$, let $\{\mu_{i,k}\}_{1 \leq k \leq N} \subset]0, +\infty[$ and $\{\vartheta_{i,j}\}_{i+1 \leq j \leq m} \subset]0, +\infty[$. The objective is to

$$\begin{aligned}
 \underset{x_1 \in \mathbb{R}^N, \dots, x_m \in \mathbb{R}^N}{\text{minimize}} \quad & \frac{p-1}{2} \sum_{i=1}^m \|x_i - \sigma_i^{-1} z_i\|^2 \\
 & + \sum_{i=1}^m \sum_{k=1}^N \mu_{i,k} \sigma_i |\langle x_i \mid e_{i,k} \rangle| \\
 & + \sum_{i=1}^{m-1} \sum_{j=i+1}^m \vartheta_{i,j} \\
 & \times \sum_{k=1}^N |\langle \sigma_i x_i - \sigma_j x_j \mid b_k \rangle| \\
 & + \sum_{i=1}^m \iota_{C_i}(\sigma_i x_i).
 \end{aligned} \tag{65}$$

To cast this problem in the format of Problem 1, let us define

$$J = \{(i, j) \in \mathbb{N}^2 \mid 1 \leq i \leq m-1, i+1 \leq j \leq m\} \tag{66}$$

and

$$\begin{aligned}
 i: \quad & J \rightarrow \{1, \dots, m(m-1)/2\} \\
 & (i, j) \mapsto m(i-1) - i(i+1)/2 + j.
 \end{aligned} \tag{67}$$

Moreover, let us set $p = m(m-1)/2 + 3$ and

$$\left\{ \begin{array}{l}
 (\forall (i, j) \in J) \\
 f_{i(i,j)}: (x_1, \dots, x_m) \\
 \quad \mapsto \vartheta_{i,j} \sum_{k=1}^N |\langle \sigma_i x_i - \sigma_j x_j \mid b_k \rangle| \\
 f_{p-2}: (x_1, \dots, x_m) \mapsto \sum_{i=1}^m \sum_{k=1}^N \mu_{i,k} \sigma_i |\langle x_i \mid e_{i,k} \rangle| \\
 f_{p-1}: (x_1, \dots, x_m) \mapsto \sum_{i=1}^m \iota_{C_i}(\sigma_i x_i) \\
 f_p: (x_1, \dots, x_m) \mapsto \frac{p-1}{2} \sum_{i=1}^m \|x_i - \sigma_i^{-1} z_i\|^2.
 \end{array} \right. \tag{68}$$

Note that, for every $k \in \{1, \dots, p-2\}$, $\text{dom } f_k = (\mathbb{R}^N)^m$ and $\text{dom } f_{p-1} = \sigma_1^{-1} C_1 \times \dots \times \sigma_m^{-1} C_m$. Hence, since the sets $(C_i)_{1 \leq i \leq m}$ are nonempty, (49) holds and Problem 7 can be solved by the Dykstra-like algorithm presented in Theorem 4, with $\mathcal{H}_1 = \dots = \mathcal{H}_m = \mathbb{R}^N$. An explicit form of the proximity operators of the functions $(f_k)_{1 \leq k \leq m(m-1)/2}$ can be deduced from Proposition 3. Indeed, for every $(i, j) \in J$, we can set in this proposition $\mathcal{H}_1 = \dots = \mathcal{H}_m = \mathcal{G} = \mathbb{R}^N$, $(\forall k \in \{1, \dots, N\})$ and $\phi_k = \vartheta_{i,j} |\cdot|$, and define the matrices $(L_\ell)_{1 \leq \ell \leq m}$ in $\mathbb{R}^{N \times N}$ as

$$(\forall \ell \in \{1, \dots, m\}) \quad L_\ell = \begin{cases} \sigma_\ell \mathbf{I}, & \text{if } \ell = i; \\ -\sigma_\ell \mathbf{I}, & \text{if } \ell = j; \\ 0 & \text{otherwise.} \end{cases} \tag{69}$$

Finally, the proximity operator of f_{p-2} can be derived from Proposition 1 combined with the separability of this function, [21, Lemmas 2.8 and 2.9], and [14, Example 4.2]. The proximity operator of f_{p-1} is provided in Example 1.

5.2.2 Numerical Experiments

Figure 3 shows the results obtained on a multispectral image of size 256×256 ($N = 256^2$) with 3 channels ($m = 3$) and pixel values in the range $[0, 255]$. These images are corrupted by white Gaussian noises with standard deviations $\sigma_1 = 11$, $\sigma_2 = 12$, and $\sigma_3 = 13$ (the corresponding SNR values are indicated in Fig. 3). On the other hand, $(b_k)_{1 \leq k \leq N}$ is the Haar orthonormal wavelet basis on 3 resolution levels and, for every $i \in \{1, \dots, m\}$, $(e_{i,k})_{1 \leq k \leq N}$ are symmlet orthonormal wavelet bases (length 6) on 3 resolution levels. The values of the regularization parameters $(\mu_{i,k})_{1 \leq i \leq 3, 1 \leq k \leq N}$ (chosen subband-adaptive by a maximum likelihood approach), and of the coupling parameters $\vartheta_{1,2}$, $\vartheta_{1,3}$, and $\vartheta_{2,3}$ are selected so as to maximize the SNR. For every $i \in \{1, \dots, m\}$, $C_i = [0, 255]^N$ models the constraint on the range of pixel values. The resulting Dykstra-like algorithm is described below.

Initialization

$$\left[\begin{array}{l}
 \sigma_1 = 11; \quad \sigma_2 = 12; \quad \sigma_3 = 13 \\
 y_{1,1,0} = \dots = y_{1,5,0} = x_{1,0} = z_1 \\
 y_{2,1,0} = \dots = y_{2,5,0} = x_{2,0} = z_2 \\
 y_{3,1,0} = \dots = y_{3,5,0} = x_{3,0} = z_3 \\
 \alpha_{1,2} = \sigma_1^2 + \sigma_2^2 \\
 \alpha_{1,3} = \sigma_1^2 + \sigma_3^2 \\
 \alpha_{2,3} = \sigma_2^2 + \sigma_3^2
 \end{array} \right.$$

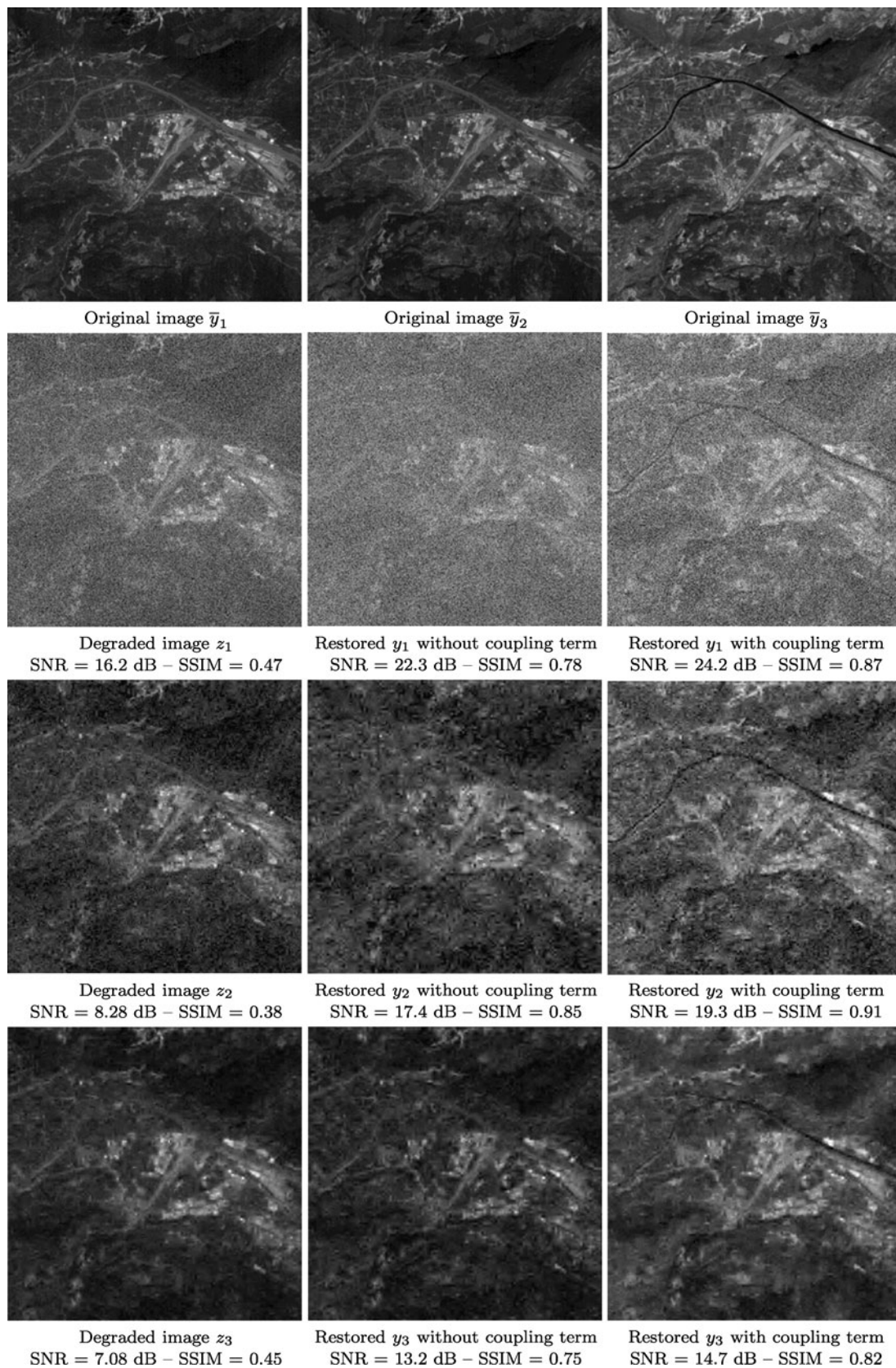


Fig. 3 Multispectral restoration

For $n = 0, 1, \dots$

$$u_{1,1,n} = y_{1,1,n} + \alpha_{1,2}^{-1} \sigma_1 \sum_{k=1}^N (\text{prox}_{\alpha_{1,2} \vartheta_{1,2} |\cdot|} (\sigma_1 y_{1,1,n} - \sigma_2 y_{2,1,n} | b_k) + \langle \sigma_1 y_{1,1,n} - \sigma_2 y_{2,1,n} | b_k \rangle) b_k$$

$$u_{2,1,n} = y_{2,1,n} - \alpha_{1,2}^{-1} \sigma_2 \sum_{k=1}^N (\text{prox}_{\alpha_{1,2} \vartheta_{1,2} |\cdot|} (\sigma_1 y_{1,1,n} - \sigma_2 y_{2,1,n} | b_k) + \langle \sigma_1 y_{1,1,n} - \sigma_2 y_{2,1,n} | b_k \rangle) b_k$$

$$u_{3,1,n} = y_{3,1,n} + \alpha_{1,3}^{-1} \sigma_1 \sum_{k=1}^N (\text{prox}_{\alpha_{1,3} \vartheta_{1,3} |\cdot|} (\sigma_1 y_{1,2,n} - \sigma_3 y_{3,2,n} | b_k) + \langle \sigma_1 y_{1,2,n} - \sigma_3 y_{3,2,n} | b_k \rangle) b_k$$

$$u_{2,2,n} = y_{2,2,n} + \alpha_{1,3}^{-1} \sigma_3 \sum_{k=1}^N (\text{prox}_{\alpha_{1,3} \vartheta_{1,3} |\cdot|} (\sigma_1 y_{1,2,n} - \sigma_3 y_{3,2,n} | b_k) + \langle \sigma_1 y_{1,2,n} - \sigma_3 y_{3,2,n} | b_k \rangle) b_k$$

$$u_{1,3,n} = y_{1,3,n} + \alpha_{2,3}^{-1} \sigma_2 \sum_{k=1}^N (\text{prox}_{\alpha_{2,3} \vartheta_{2,3} |\cdot|} (\sigma_2 y_{2,3,n} - \sigma_3 y_{3,3,n} | b_k) + \langle \sigma_2 y_{2,3,n} - \sigma_3 y_{3,3,n} | b_k \rangle) b_k$$

$$u_{3,3,n} = y_{3,3,n} - \alpha_{2,3}^{-1} \sigma_3 \sum_{k=1}^N (\text{prox}_{\alpha_{2,3} \vartheta_{2,3} |\cdot|} (\sigma_2 y_{2,3,n} - \sigma_3 y_{3,3,n} | b_k) + \langle \sigma_2 y_{2,3,n} - \sigma_3 y_{3,3,n} | b_k \rangle) b_k$$

$$u_{1,4,n} = \sum_{k=1}^N (\text{prox}_{\mu_{1,k} \sigma_{1|\cdot|}} \langle y_{1,4,n} | e_{1,k} \rangle) e_{1,k}$$

$$u_{2,4,n} = \sum_{k=1}^N (\text{prox}_{\mu_{2,k} \sigma_{2|\cdot|}} \langle y_{2,4,n} | e_{2,k} \rangle) e_{2,k}$$

$$u_{3,4,n} = \sum_{k=1}^N (\text{prox}_{\mu_{3,k} \sigma_{3|\cdot|}} \langle y_{3,4,n} | e_{3,k} \rangle) e_{3,k}$$

$$u_{1,5,n} = P_{C_1}(\sigma_1 y_{1,5,n})$$

$$u_{2,5,n} = P_{C_2}(\sigma_2 y_{2,5,n})$$

$$u_{3,5,n} = P_{C_3}(\sigma_3 y_{3,5,n})$$

$$x_{1,n+1} = (u_{1,1,n} + u_{1,2,n} + u_{1,3,n} + u_{1,4,n} + u_{1,5,n})/5$$

$$x_{2,n+1} = (u_{2,1,n} + u_{2,2,n} + u_{2,3,n} + u_{2,4,n} + u_{2,5,n})/5$$

$$x_{3,n+1} = (u_{3,1,n} + u_{3,2,n} + u_{3,3,n} + u_{3,4,n} + u_{3,5,n})/5$$

For $j = 1, \dots, 5$

$$\begin{cases} y_{1,j,n+1} = x_{1,n+1} + y_{1,j,n} - u_{1,j,n} \\ y_{2,j,n+1} = x_{2,n+1} + y_{2,j,n} - u_{2,j,n} \\ y_{3,j,n+1} = x_{3,n+1} + y_{3,j,n} - u_{3,j,n} \end{cases}$$

It can be observed from the images displayed on the second and third columns of Fig. 3 that the introduction of the coupling term has a significant influence on denoising performance. Moreover, in our experiments, we observed that better results were obtained when different bases $(b_k)_{1 \leq k \leq N}, (e_{1,k})_{1 \leq k \leq N}, \dots, (e_{m,k})_{1 \leq k \leq N}$ were employed.

It turns out that, in this particular problem, an alternative solution method is PPXA (see Theorem 3) applied to the minimization of the sum of the $m(m-1)/2 + 1$ functions f_1, f_2, \dots, f_{p-2} , and $f_{p-1} + f_p$ defined in (68). The proximity operator of the latter is given by

[21, Lemma 2.6(i)]. Indeed, the qualification condition (see (48)) is satisfied since $\text{dom } f_1 = \dots = \text{dom } f_{p-2} = (\mathbb{R}^N)^m$ and, $(\forall i \in \{1, \dots, m\}) \text{int } C_i =]0, 255[\neq \emptyset$. The choice of the PPXA parameters has been optimized empirically for speed of convergence and set to $\lambda_n \equiv 1.3, \gamma = 1$, and $\omega_1 = \dots = \omega_{p-1} = 1/(p-1)$. In Fig. 4, $\|x_n - x_\infty\|/\|x_0 - x_\infty\|$ is plotted as a function of computation time, where $(x_n)_{n \in \mathbb{N}} = ((x_{1,n}, x_{2,n}, x_{3,n}))_{n \in \mathbb{N}}$ is the sequence generated by an algorithm and x_∞ is the unique solution to Problem 7. In our experiments, 500 iterations were used to produce this solution.

5.3 Structure-Texture Image Decomposition

An important problem in image processing is to decompose an image into elementary structures. In the context of denoising, this decomposition was investigated in [37] with a total variation potential. In [31], a different potential was used to better penalize strongly oscillating components. The resulting variational problem is not straightforward. Numerical methods were proposed in [5, 41] and experiments were performed for image denoising and analysis problems based on a geometry-texture decomposition. Another challenging problem is to extract meaningful components from a blurred and noise-corrupted image. In the presence of additive Gaussian noise, a decomposition into geometry and texture components is proposed in [2, 23]. The method developed in the present paper, will make it possible to consider general (not necessarily additive and Gaussian) noise models and arbitrary linear degradation operators. We consider a simple geometry-texture decomposition from a degraded observation.

5.3.1 Problem Formulation

In this experiment, the observed image $z \in \mathbb{R}^N$ is obtained by multiplying the original image $\bar{x} \in \mathbb{R}^N$ with a matrix $T \in \mathbb{R}^{N \times N}$, which models a blur, and corrupting $T\bar{x}$ by a Poisson noise with scaling parameter α . It is assumed that

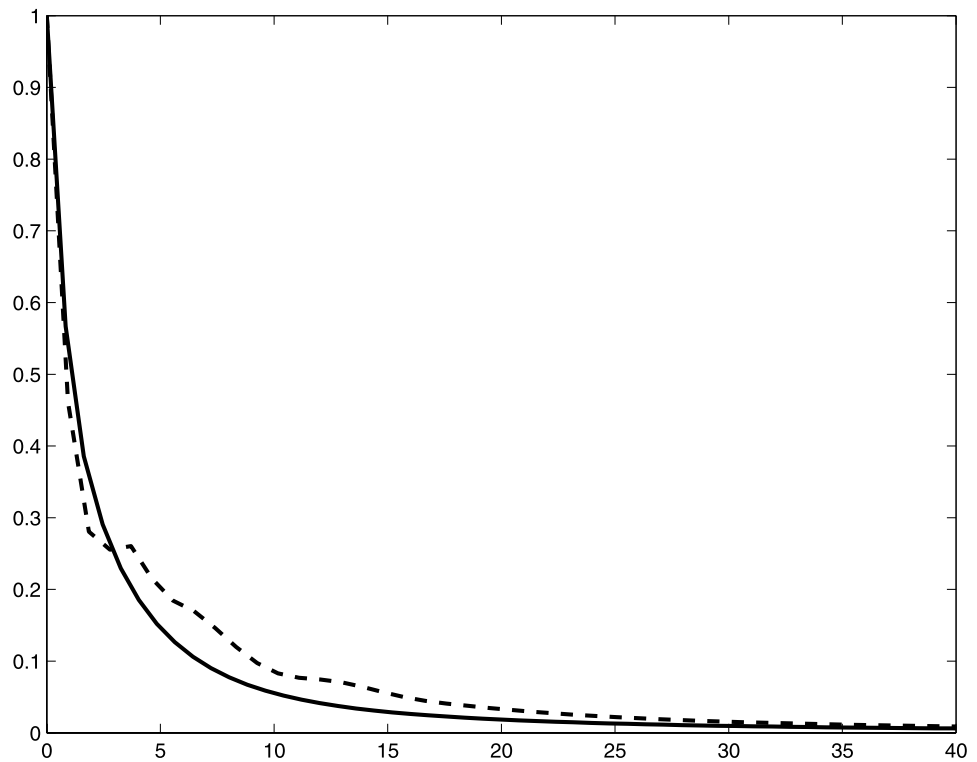
$$T \text{ has its entries in } [0, +\infty[\text{ and each of its rows is nonzero.} \tag{70}$$

The inverse problem we address is to obtain the decomposition of \bar{x} into the sum of a geometry and a texture component, say

$$\bar{x} = R_1(\bar{x}_1) + R_2(\bar{x}_2), \tag{71}$$

where $R_1 : \mathbb{R}^{N_1} \mapsto \mathbb{R}^N$ and $R_2 : \mathbb{R}^{N_2} \mapsto \mathbb{R}^N$ are known operators. The vectors $\bar{x}_1 \in \mathbb{R}^{N_1}$ and $\bar{x}_2 \in \mathbb{R}^{N_2}$ to be estimated parameterize, respectively, the geometry and the texture components.

Fig. 4 Problem 7: Convergence profiles of the Dykstra-like algorithm (*solid line*) and of PPXA (*dashed line*) versus computation time in seconds



We consider a simple instance of (71) involving a linear mixture: $N_1 = N$, $R_1: x_1 \mapsto x_1$, and $R_2: x_2 \mapsto F^T x_2$, where $F^T \in \mathbb{R}^{N \times K}$ is a linear tight frame synthesis operator. In other words, the information regarding the texture component pertains to the coefficients \bar{x}_2 of its decomposition in the frame. The tightness condition implies that

$$F^T F = \nu \text{Id}, \quad \text{for some } \nu \in]0, +\infty[. \tag{72}$$

Thus, the original image is decomposed as $\bar{x} = \bar{x}_1 + F^T \bar{x}_2$. It is known a priori that $\bar{x} \in C_1 \cap C_2$, where

$$C_1 = [0, 255]^N \tag{73}$$

models the constraint on the range of pixel values, and

$$C_2 = \left\{ x \in \mathbb{R}^N \mid \widehat{x} = (\eta_k)_{1 \leq k \leq N}, \sum_{k \in \mathbb{I}} |\eta_k|^2 \leq \delta \right\}, \tag{74}$$

for some $\delta \in]0, +\infty[$, models an energy constraint on the 2-D DFT \widehat{x} of the original image in some low frequency band $\mathbb{I} \subset \{1, \dots, N\}$. In addition, to limit the total variation [11] of the geometrical component, the potential $x \mapsto \psi(Hx, Vx)$ is used, with

$$\begin{aligned} \psi: & ((\eta_k)_{1 \leq k \leq N}, (\zeta_k)_{1 \leq k \leq N}) \\ & \mapsto \chi \sum_{k=1}^N \sqrt{|\eta_k|^2 + |\zeta_k|^2}, \end{aligned} \tag{75}$$

where $H \in \mathbb{R}^{N \times N}$ and $V \in \mathbb{R}^{N \times N}$ are matrix representations of the horizontal and vertical discrete differentiation operators, respectively, and where $\chi \in]0, +\infty[$. Furthermore, to promote sparsity in the frame of the texture component of the image, the potential

$$h: (\eta_k)_{1 \leq k \leq K} \mapsto \sum_{k=1}^K \tau_k |\eta_k| \tag{76}$$

is introduced, where $\{\tau_k\}_{1 \leq k \leq K} \subset]0, +\infty[$. Finally, as a data fidelity term well adapted to Poisson noise, we employ the generalized Kullback-Leibler divergence with a scaling parameter $\alpha \in]0, +\infty[$. Upon setting $z = (\zeta_k)_{1 \leq k \leq N}$, this leads to the function

$$g: (\xi_k)_{1 \leq k \leq N} \mapsto \sum_{k=1}^N \phi_k(\xi_k), \tag{77}$$

where, for every $k \in \{1, \dots, K\}$,

$$\begin{aligned} \phi_k: & \mathbb{R} \rightarrow]-\infty, +\infty] \\ \xi \mapsto & \begin{cases} -\zeta_k \ln(\xi) + \alpha \xi, & \text{if } \zeta_k \geq 0 \text{ and } \xi > 0; \\ 0, & \text{if } \xi = 0; \\ +\infty, & \text{otherwise.} \end{cases} \end{aligned} \tag{78}$$

Altogether, the variational problem is to

$$\begin{aligned} & \underset{\substack{x_1 \in \mathbb{R}^N, x_2 \in \mathbb{R}^K \\ x_1 + F^\top x_2 \in C_1 \\ x_1 + F^\top x_2 \in C_2}}{\text{minimize}} & \psi(Hx_1, Vx_1) + h(x_2) + g(Tx_1 + TF^\top x_2). \end{aligned} \tag{79}$$

This problem is a particular case of (2) with $m = 2, p = 4$, and

$$\begin{cases} f_1: (x_1, x_2) \mapsto \psi(Hx_1, Vx_1) + h(x_2), \\ f_2: (x_1, x_2) \mapsto g(Tx_1 + TF^\top x_2), \\ f_3: (x_1, x_2) \mapsto \iota_{C_1}(x_1 + F^\top x_2), \\ f_4: (x_1, x_2) \mapsto \iota_{C_2}(x_1 + F^\top x_2). \end{cases} \tag{80}$$

However, since the operators $(\text{prox}_{f_i})_{1 \leq i \leq 4}$ are not easily implementable, we cannot apply directly Theorems 1, 2, or 3. To circumvent this difficulty, a strategy is to decompose (79) into an equivalent problem by introducing auxiliary variables.

A first equivalent problem to (79) is

$$\begin{aligned} & \underset{\substack{x_1, x_2, x_3, x_4, x_5, x_6 \\ x_3 = x_1 + F^\top x_2 \\ x_3 \in C_1 \cap C_2 \\ x_4 = Tx_3 \\ x_5 = Hx_1, x_6 = Vx_1}}{\text{minimize}} & \psi(x_5, x_6) + h(x_2) + g(x_4), \end{aligned} \tag{81}$$

where we have introduced the auxiliary variables $(x_3, x_4, x_5, x_6) \in \mathbb{R}^N \oplus \mathbb{R}^N \oplus \mathbb{R}^N \oplus \mathbb{R}^N$. Problem (81) is a particular case of (2) with $m = 6, p = 3$, and

$$\begin{cases} f_1: (x_1, \dots, x_6) \mapsto h(x_2) + \iota_{C_1}(x_3) \\ \qquad \qquad \qquad + g(x_4) + \psi(x_5, x_6), \\ f_2: (x_1, \dots, x_6) \mapsto \iota_{C_2}(x_3), \\ f_3: (x_1, \dots, x_6) \mapsto \iota_{\{0\}}(x_1 + F^\top x_2 - x_3) \\ \qquad \qquad \qquad + \iota_{\{0\}}(Tx_3 - x_4) \\ \qquad \qquad \qquad + \iota_{\{0\}}(Hx_1 - x_5) \\ \qquad \qquad \qquad + \iota_{\{0\}}(Vx_1 - x_6). \end{cases} \tag{82}$$

In this formulation, the role of f_3 is to impose the constraints $x_1 + F^\top x_2 = x_3, Tx_3 = x_4, Hx_1 = x_5$, and $Vx_1 = x_6$. As seen in Example 1, $\text{prox}_{\iota_{C_1}} = P_{C_1}$ and $\text{prox}_{\iota_{C_2}} = P_{C_2}$. On the other hand, the proximity operators of ψ, h , and g can be obtained from [19, Proposition 2.8(i)], [21, Example 2.16], and [17, Example 30], respectively. In turn, since f_1 is separable, its proximity operator follows straightforwardly componentwise. Now set

$$L_1 = \begin{bmatrix} I & F^\top & -I & [0] & [0] & [0] \\ [0] & [0] & T & -I & [0] & [0] \\ H & [0] & [0] & [0] & -I & [0] \\ V & [0] & [0] & [0] & [0] & -I \end{bmatrix}. \tag{83}$$

It follows from (82) and (83) that $f_3 = \iota_{\ker L_1}$, where $\ker L_1 = \{x \in \mathcal{H} \mid L_1 x = 0\}$. Hence, by Example 1 and [24, Chap. 8],

$$\text{prox}_{f_3} = P_{\ker L_1} = I - L_1^\top (L_1 L_1^\top)^{-1} L_1. \tag{84}$$

Under the assumption that the matrices T, H , and V are block-circulant with circulant blocks, they are diagonalized by the 2-D DFT. Hence, combining (84), (83), and (72) we deduce that prox_{f_3} is computable explicitly. On the other hand, it follows from (82), (76), (73), (77), (75), (74), and (83) that

$$\begin{cases} \text{ri dom } f_1 = \mathbb{R}^N \times \mathbb{R}^K \times \text{int } C_1 \\ \qquad \qquad \times]0, +\infty[^N \times \mathbb{R}^N \times \mathbb{R}^N \\ \text{ri dom } f_2 = \mathbb{R}^N \times \mathbb{R}^K \times \text{int } C_2 \\ \qquad \qquad \times \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N \\ \text{ri dom } f_3 = \ker L_1. \end{cases} \tag{85}$$

Hence, qualification condition (48) reduces to

$$(\exists(x_1, \dots, x_6) \in \ker L_1) \begin{cases} x_3 \in \text{int } C_1 \cap \text{int } C_2, \\ x_4 \in]0, +\infty[^N, \end{cases} \tag{86}$$

which is equivalent to

$$\begin{aligned} & (\exists(x_1, \dots, x_6) \in \mathbb{R}^N \times \mathbb{R}^K \times \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N) \\ & \begin{cases} x_1 + F^\top x_2 = x_3 \in \text{int } C_1 \cap \text{int } C_2, \\ Tx_3 = x_4 \in]0, +\infty[^N \\ Hx_1 = x_5 \\ Vx_1 = x_6. \end{cases} \end{aligned} \tag{87}$$

This condition is satisfied if

$$T(\text{int}(C_1 \cap C_2)) \cap]0, +\infty[^N \neq \emptyset. \tag{88}$$

Indeed, let $y \in T(\text{int}(C_1 \cap C_2)) \cap]0, +\infty[^N$. Then there exists $x \in \text{int}(C_1 \cap C_2)$ such that $Tx = y \in]0, +\infty[^N$. Hence, for every $x_2 \in \mathbb{R}^K$, if we set $x_3 = x, x_4 = y = Tx_3, x_1 = x_3 - F^\top x_2, x_5 = Hx_1$, and $Vx_1 = x_6$, (87) is seen to hold. Since (73) and (74) yield $\text{int}(C_1 \cap C_2) \neq \emptyset$, we deduce from (70) that (88) (and hence (48)) is satisfied. Thus, (81) can be solved by PPXA (see Theorem 3 and Remark 3).

Another equivalent formulation of (79) is

$$\begin{aligned} & \underset{\substack{x_1, x_2, x_3, x_4, x_5, x_6, x_7 \\ x_3 = x_1 + F^\top x_2 \\ x_4 = Tx_3 \\ x_5 = Hx_1, x_6 = Vx_1 \\ x_7 = x_3 \\ x_3 \in C_1, x_7 \in C_2}}{\text{minimize}} & \psi(x_5, x_6) + h(x_2) + g(x_4), \end{aligned} \tag{89}$$

where the additional auxiliary variable $x_7 \in \mathbb{R}^N$ has been introduced. Problem (89) is a particular case of (2) with $m =$

7, $p = 2$, and

$$\begin{cases} f_1: (x_1, \dots, x_7) \mapsto h(x_2) + \iota_{C_1}(x_3) \\ \quad \quad \quad + g(x_4) + \psi(x_5, x_6) \\ \quad \quad \quad + \iota_{C_2}(x_7) \\ f_2: (x_1, \dots, x_7) \mapsto \iota_{\{0\}}(x_1 + F^\top x_2 - x_3) \\ \quad \quad \quad + \iota_{\{0\}}(Tx_3 - x_4) \\ \quad \quad \quad + \iota_{\{0\}}(Hx_1 - x_5) \\ \quad \quad \quad + \iota_{\{0\}}(Vx_1 - x_6) + \iota_{\{0\}}(x_3 - x_7). \end{cases} \quad (90)$$

As previously observed, since the proximity operators of ψ , h , g , ι_{C_1} , and ι_{C_2} are easily computable, so is prox_{f_i} . Furthermore, if we set

$$L_2 = \begin{bmatrix} I & F^\top & -I & [0] & [0] & [0] & [0] \\ [0] & [0] & T & -I & [0] & [0] & [0] \\ H & [0] & [0] & [0] & -I & [0] & [0] \\ V & [0] & [0] & [0] & [0] & -I & [0] \\ [0] & [0] & I & [0] & [0] & [0] & -I \end{bmatrix}, \quad (91)$$

it can be deduced from (90) that the proximity operator of $f_2 = \iota_{\ker L_2}$ can be computed like that of $\iota_{\ker L_1}$. We derive from (90), (76), (73), (77), (75), (74), and (91) that

$$\begin{cases} \text{ri dom } f_1 = \mathbb{R}^N \times \mathbb{R}^K \times \text{int } C_1 \\ \quad \quad \quad \times]0, +\infty[^N \times \mathbb{R}^N \times \mathbb{R}^N \times \text{int } C_2 \\ \text{ri dom } f_2 = \ker L_2. \end{cases} \quad (92)$$

Hence, arguing as above, (45) reduces to (88), which is seen to be satisfied. This shows that (89) can be solved by the Douglas-Rachford algorithm (see Theorem 2 and Remark 22).

5.3.2 Numerical Experiments

Figure 5 shows the results of the decomposition into geometry and texture components of an electron microscopy image of size 512×512 ($N = 512^2$) degraded by a Gaussian blur of size 5×5 and Poisson noise with scaling parameter $\alpha = 0.5$. The parameter χ of (75) and the parameters $(\tau_k)_{1 \leq k \leq K}$ of (76) are selected so as to maximize the SNR. The matrix F is a tight frame version of the dual-tree transform proposed in [13] using symmlet of length 6 applied over 3 resolution levels ($\nu = 2$ and $K = 2N$). The same discrete gradient matrices H and V as in [5] are used. We aim at comparing the PPXA and Douglas-Rachford algorithms in the image decomposition problem under consideration. In both algorithms we set $\lambda_n \equiv 1$.

In this context, setting $\omega_1 = \omega_2 = \omega_3 = 1/3$, PPXA assumes the following form.

Initialization

$$\begin{cases} \gamma = 100 \\ (y_{1,1,0}, \dots, y_{6,1,0}) = (z, F^\top z, z, z, z, z) \\ (y_{1,2,0}, \dots, y_{6,2,0}) = (z, F^\top z, z, z, z, z) \\ (y_{1,3,0}, \dots, y_{6,3,0}) = (z, F^\top z, z, z, z, z) \\ \text{For } i = 1, \dots, 6 \\ \lfloor x_{i,0} = (y_{i,1,0} + y_{i,2,0} + y_{i,3,0})/3 \end{cases}$$

For $n = 0, 1, \dots$

$$\begin{cases} u_{1,1,n} = y_{1,1,n} \\ u_{2,1,n} = \text{prox}_{3\gamma h}(y_{2,1,n}) \\ u_{3,1,n} = P_{C_1}(y_{3,1,n}) \\ u_{4,1,n} = \text{prox}_{3\gamma g}(y_{4,1,n}) \\ (u_{5,1,n}, u_{6,1,n}) = \text{prox}_{3\gamma \psi}(y_{5,1,n}, y_{6,1,n}) \\ (u_{1,2,n}, u_{2,2,n}) = (y_{1,2,n}, y_{2,2,n}) \\ u_{3,2,n} = P_{C_2}(y_{3,2,n}) \\ (u_{4,2,n}, u_{5,2,n}, u_{6,2,n}) = (y_{4,2,n}, y_{5,2,n}, y_{6,2,n}) \\ (u_{1,3,n}, \dots, u_{6,3,n}) = P_{\ker L_1}(y_{1,3,n}, \dots, y_{6,3,n}) \\ \text{For } i = 1, \dots, 6 \\ \begin{cases} s_{i,n} = (1/3) \sum_{k=1}^3 u_{i,k,n} \\ y_{i,1,n+1} = y_{i,1,n} + 2s_{i,n} - x_{i,n} - u_{i,1,n} \\ y_{i,2,n+1} = y_{i,2,n} + 2s_{i,n} - x_{i,n} - u_{i,2,n} \\ y_{i,3,n+1} = y_{i,3,n} + 2s_{i,n} - x_{i,n} - u_{i,3,n} \\ x_{i,n+1} = x_{i,n} + s_{i,n} - x_{i,n} \end{cases} \end{cases}$$

On the other hand, the Douglas-Rachford algorithm reduces to the following.

Initialization

$$\begin{cases} \gamma = 100 \\ (x_{1,0}, \dots, x_{7,0}) = (z, F^\top z, z, z, z, z, z) \end{cases}$$

For $n = 0, 1, \dots$

$$\begin{cases} y_{1,n} = x_{1,n} \\ y_{2,n} = \text{prox}_{\gamma h}(x_{2,n}) \\ y_{3,n} = P_{C_1}(x_{3,n}) \\ y_{4,n} = \text{prox}_{\gamma g}(x_{4,n}) \\ (y_{5,n}, y_{6,n}) = \text{prox}_{\gamma \psi}(x_{5,n}, x_{6,n}) \\ y_{7,n} = P_{C_2}(x_{7,n}) \\ (u_{1,n}, \dots, u_{7,n}) \\ \quad = P_{\ker L_2}(2(y_{1,n}, \dots, y_{7,n}) - (x_{1,n}, \dots, x_{7,n})) \\ \text{For } i = 1, \dots, 7 \\ \lfloor x_{i,n+1} = x_{i,n} + u_{i,n} - y_{i,n} \end{cases}$$

In Fig. 6, the value of $\|y_n - y_\infty\|/\|y_0 - y_\infty\|$ for the sequence $(y_n)_{n \in \mathbb{N}} = ((y_{1,n}, \dots, y_{7,n}))_{n \in \mathbb{N}}$ of Theorem 2 and $\|x_n - x_\infty\|/\|x_0 - x_\infty\|$ for the sequence $(x_n)_{n \in \mathbb{N}} = ((x_{1,n}, \dots, x_{6,n}))_{n \in \mathbb{N}}$ of Theorem 3 (where y_∞ and x_∞ denote the respective limits) are plotted as a function of the

Fig. 5 Decomposition and restoration results

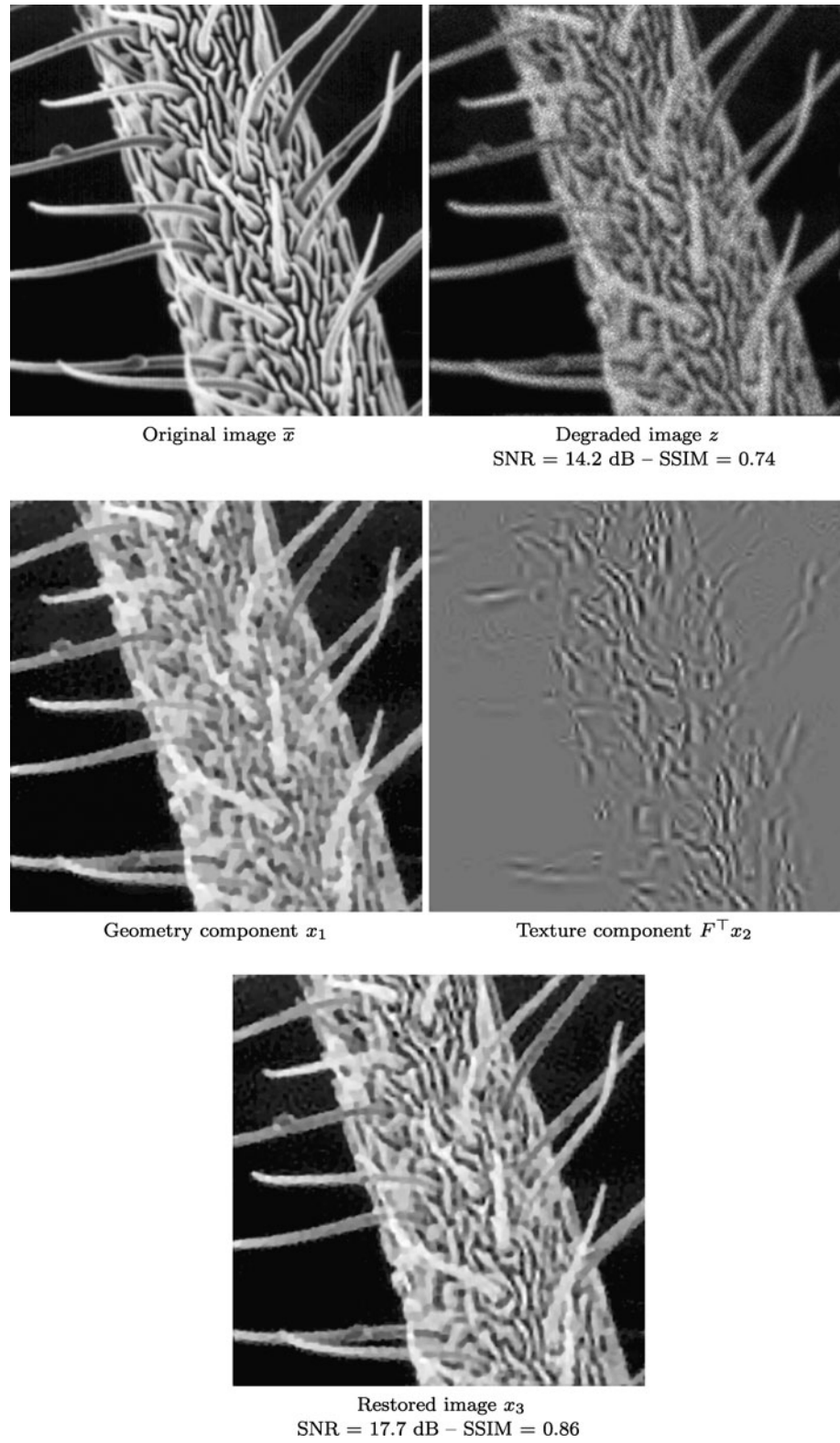
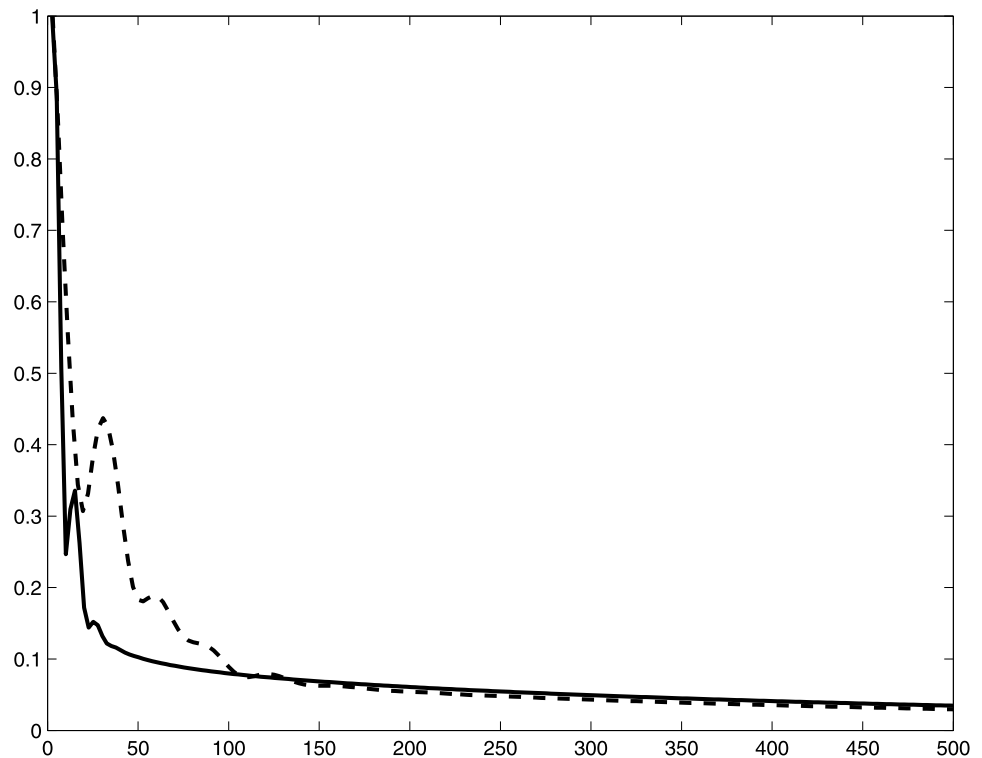


Fig. 6 Convergence profiles of the Douglas-Rachford algorithm (solid line) and PPXA (dashed line) versus computation time in seconds



computation time in seconds. In our experiments, 1000 iterations were used to produce a solution.

6 Conclusion

In this paper, the proximal formalism has been applied to multicomponent signal/image processing. Expressions of new proximity operators in product spaces have been derived. The proposed multicomponent framework has been illustrated through three different applications: stereoscopy, multispectral imagery, and decomposition into geometry and texture components. Another field of application in which these techniques could be useful is the processing of color images. The proposed proximal formalism can also be used to derive algorithms for complex signal and image processing by regarding a complex signal as a signal with $m = 2$ real components, namely its real and imaginary parts.

References

1. Afonso, M.V., Bioucas-Dias, J.M., Figueiredo, M.A.T.: An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems. *IEEE Trans. Image Process.* (2010). doi:10.1109/TIP.2010.2076294
2. Anthoine, S., Pierpaoli, E., Daubechies, I.: Deux méthodes de déconvolution et séparation simultanées; application à la reconstruction des amas de galaxies. *Trait. Signal* **23**, 439–447 (2006)
3. Antoniadis, A., Fan, J.: Regularization of wavelet approximations. *J. Am. Stat. Assoc.* **96**, 939–967 (2001)
4. Aujol, J.-F., Chambolle, A.: Dual norms and image decomposition models. *Int. J. Comput. Vis.* **63**, 85–104 (2005)
5. Aujol, J.-F., Aubert, G., Blanc-Féraud, L., Chambolle, A.: Image decomposition into a bounded variation component and an oscillating component. *J. Math. Imaging Vis.* **22**, 71–88 (2005)
6. Aujol, J.-F., Gilboa, G., Chan, T., Osher, S.J.: Structure-texture image decomposition—Modeling, algorithms, and parameter selection. *Int. J. Comput. Vis.* **67**, 111–136 (2006)
7. Aujol, J.-F., Kang, S.H.: Color image decomposition and restoration. *J. Vis. Commun. Image Represent.* **17**, 916–928 (2006)
8. Bauschke, H.H., Combettes, P.L.: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York (2011)
9. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2**, 183–202 (2009)
10. Briceño-Arias, L.M., Combettes, P.L.: Convex variational formulation with smooth coupling for multicomponent signal decomposition and recovery. *Numer. Math. Theory Methods Appl.* **2**, 485–508 (2009)
11. Chambolle, A.: An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.* **20**, 89–97 (2004)
12. Chan, T.F., Esedoglu, S., Park, F.E.: Image decomposition combining staircase reduction and texture extraction. *J. Vis. Commun. Image Represent.* **18**, 464–486 (2007)
13. Chaux, C., Duval, L., Pesquet, J.-C.: Image analysis using a dual-tree M -band wavelet transform. *IEEE Trans. Image Process.* **15**, 2397–2412 (2006)
14. Chaux, C., Combettes, P.L., Pesquet, J.-C., Wajs, V.R.: A variational formulation for frame-based inverse problems. *Inverse Probl.* **23**, 1495–1518 (2007)
15. Chaux, C., Benazza-Benyahia, A., Pesquet, J.-C., Duval, L.: Wavelet transform for the denoising of multivariate images. In: Collet, C., Chanussot, J., Chehdi, K. (eds.) *Multivariate Image Processing*, pp. 203–237. Wiley, New York (2010)

16. Combettes, P.L.: Iterative construction of the resolvent of a sum of maximal monotone operators. *J. Convex Anal.* **16**, 727–748 (2009)
17. Combettes, P.L., Pesquet, J.-C.: A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE J. Selected Topics. Signal Process.* **1**, 564–574 (2007)
18. Combettes, P.L., Pesquet, J.-C.: Proximal thresholding algorithm for minimization over orthonormal bases. *SIAM J. Optim.* **18**, 1351–1376 (2007)
19. Combettes, P.L., Pesquet, J.-C.: A proximal decomposition method for solving convex variational inverse problems. *Inverse Probl.* **24**, 065014 (2008)
20. Combettes, P.L., Pesquet, J.-C.: Proximal splitting methods in signal processing. In: Bauschke, H.H., et al. (eds.) *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. Springer, New York (2010)
21. Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.* **4**, 1168–1200 (2005)
22. Daubechies, I., Defrise, M., De Mol, C.: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math.* **57**, 1413–1457 (2004)
23. Daubechies, I., Teschke, G.: Variational image restoration by means of wavelets: simultaneous decomposition, deblurring and denoising. *Appl. Comput. Harmon. Anal.* **19**, 1–16 (2005)
24. Deutsch, F.: *Best Approximation in Inner Product Spaces*. Springer, New York (2001)
25. Elad, M., Starck, J.-L., Donoho, D., Querre, P.: Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA). *Appl. Comput. Harmon. Anal.* **19**, 340–358 (2005)
26. Goldburg, M., Marks, R.J. II: Signal synthesis in the presence of an inconsistent set of constraints. *IEEE Trans. Circuits Syst.* **32**, 647–663 (1985)
27. Huang, Y., Ng, M.K., Wen, Y.-W.: A fast total variation minimization method for image restoration. *Multiscale Model. Simul.* **7**, 774–795 (2008)
28. Hunt, B.R., Kübler, O.: Karhunen-Loève multispectral image restoration, part I: theory. *IEEE Trans. Acoust. Speech Signal Process.* **32**, 592–600 (1984)
29. Kang, M.: Generalized multichannel image deconvolution approach and its applications. *Opt. Eng.* **37**, 2953–2964 (1998)
30. Katsaggelos, A., Lay, K., Galatsanos, N.: A general framework for frequency domain multi-channel signal processing. *IEEE Trans. Image Process.* **2**, 417–420 (1993)
31. Meyer, Y.: *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*. AMS, Providence (2001)
32. Miled, W., Pesquet, J.-C., Parent, M.: A convex optimization approach for depth estimation under illumination variation. *IEEE Trans. Image Process.* **18**, 813–830 (2009)
33. Moreau, J.-J.: Proximité et dualité dans un espace Hilbertien. *Bull. Soc. Math. Fr.* **93**, 273–299 (1965)
34. Nesterov, Yu.: A method of solving a convex programming problem with convergence rate $O(1/k^2)$. *Sov. Math. Dokl.* **27**, 372–376 (1983)
35. Nesterov, Yu.: Primal-dual subgradient methods for convex problems. *Math. Program.* **120**, 221–259 (2009)
36. Pedone, M., Heikkilä, J.: Blur and contrast invariant fast stereo matching. *Lect. Notes Comput. Sci.* **5259**, 883–890 (2008)
37. Rudin, L.I., Osher, S.J., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Physica D* **60**, 259–268 (1992)
38. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comput. Vis.* **47**, 7–42 (2002)
39. Svaiter, B.F.: On weak convergence of the Douglas-Rachford method. *SIAM J. Control Optim.* (to appear)
40. Tschumperlé, D., Deriche, R.: Diffusion PDEs on vector-valued images. *IEEE Signal Process. Mag.* **19**, 16–25 (2002)
41. Vese, L.A., Osher, S.J.: Modeling textures with total variation minimization and oscillating patterns in image processing. *J. Sci. Comput.* **19**, 553–572 (2003)
42. Vese, L.A., Osher, S.J.: Image denoising and decomposition with total variation minimization and oscillatory functions. *J. Math. Imaging Vis.* **20**, 7–18 (2004)
43. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004)
44. Wang, Y., Yang, J., Yin, W., Zhang, Y.: A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. Imaging Sci.* **1**, 248–272 (2008)
45. Weiss, P., Aubert, G., Blanc-Féraud, L.: Efficient schemes for total variation minimization under constraints in image processing. *SIAM J. Sci. Comput.* **31**, 2047–2080 (2009)
46. Wen, Y.-W., Ng, M.K., Ching, W.-K.: Iterative algorithms based on decoupling of deblurring and denoising for image restoration. *SIAM J. Sci. Comput.* **30**, 2655–2674 (2008)
47. Zălinescu, C.: *Convex Analysis in General Vector Spaces*. World Scientific, River Edge (2002)



L.M. Briceño-Arias received a Masters degree in Transport Engineering from the Universidad de Chile, Santiago, Chile in 2006, and a masters degree in mathematics and applications from the Université Pierre et Marie Curie–Paris 6, Paris, France in 2007. He is currently a Ph.D. student in mathematics at the Université Pierre et Marie Curie–Paris 6, Paris, France.



P.L. Combettes is a Professor with the Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie–Paris 6, Paris, France. Dr. Combettes is recipient of the IEEE Signal Processing Society Paper Award and is a Fellow of the IEEE. His research interests lie mostly in nonlinear analysis and signal recovery.



J.-C. Pesquet received the engineering degree from Supélec, Gif-sur-Yvette, France, in 1987, the Ph.D. degree from the Université Paris-Sud (XI), Paris, France, in 1990, and the Habilitation à Diriger des Recherches from the Université Paris-Sud in 1999. From 1991 to 1999, he was a Maître de Conférences at the Université Paris-Sud, and a Research Scientist at the Laboratoire des Signaux et Systèmes, Centre National de la Recherche Scientifique (CNRS), Gif sur Yvette.

He is currently a Professor with Université de Paris-Est Marne-la-Vallée, France and a Research Scientist at the Laboratoire d'Informatique of the university (UMR-CNRS 8049).



N. Pustelnik received the Ingénieur degree from the Institut des Sciences de l'Ingénieur de Toulon et du Var (ISITV), France and the Master degree in telecommunication and teledetection from the Université de Toulon et du Var, France, in 2007. In 2010, she received the Ph.D. degree in signal and image processing from the Université Paris-Est, France on the topic of inverse problems resolution using proximal algorithms. She is currently with the Laboratoire Automatique, Productique et Signal (LAPS), Université de Bordeaux, France.