

# HYBRID REGULARIZATION FOR DATA RESTORATION IN THE PRESENCE OF POISSON NOISE

*Nelly Pustelnik, Caroline Chaux and Jean-Christophe Pesquet*

Université Paris-Est  
Institut Gaspard Monge, UMR CNRS 8049  
77454 Marne-la-Vallée Cedex, France  
e-mail: {nelly.pustelnik,caroline.chaux,jean-christophe.pesquet}@univ-paris-est.fr

## ABSTRACT

During the last five years, several convex optimization algorithms have been proposed for solving inverse problems. Most of the time, they allow us to minimize a criterion composed of two terms one of which permits to “stabilize” the solution. Different choices are possible for the so-called regularization term, which plays a prominent role for solving ill-posed problems. While a total variation regularization introduces staircase effects, a wavelet regularization may bring other kinds of visual artefacts. A compromise can be envisaged combining these regularization functions. In the context of Poisson data, we propose in this paper an algorithm to achieve the minimization of the associated (possibly constrained) convex optimization problem.

## 1. INTRODUCTION

Different algorithms in convex optimization have been proposed to minimize regularized criteria used to provide solutions to ill-posed problems. In this paper, we are more specifically interested in restoring degraded images when the noise is Poisson distributed.

A way to take into account Poisson noise consists of applying a pre-processing on the data (like the Anscombe [1] or the Haar-Fisz [2] transforms) in order to stabilize the variance and then to apply standard restoration tools [3, 4].

One of the first restoration method aiming at maximizing the Poisson likelihood is the Richardson-Lucy algorithm [5, 6]. However, the main drawback of this method is that it does not allow us to incorporate information about the target solution, except the positivity constraint in an implicit manner. An improved form of this algorithm involving a total variation [7] regularization was recently proposed in [8]. Due to the equivalence between the maximization of the Poisson likelihood and the minimization of a Kullback-Leibler divergence, other forms of multiplicative iterative algorithms have been proposed in order to take into account specific forms of regularizations [9].

However, during the last decade, much interest has been gained in introducing a priori information about the target image in a transformed domain. Actually, wavelets are often used due to their ability to provide sparse representations for many classes of regular signals. In this respect, redundant frames (overcomplete transforms) constitute more flexible tools than orthonormal bases for building linear representations of images. A number of recent works [10] have emphasized the interest in using specific tight frames for performing geometrical analyses of images.

Existing works for solving inverse problems in a frame are often restricted to noise with log-likelihood having a Lip-

schitz continuous gradient [10]. To alleviate this restriction, the authors in [11, 12] considered algorithms mixing forward-backward [10] and Douglas-Rachford algorithms. First, Dupé et al. [11] investigated an elegant adaptation of the Anscombe approach and then, in [12] a quadratic extension technique was introduced.

One of the drawbacks of the approaches based on wavelet representations is that they may introduce visual artefacts, e.g. some lack of regularity in homogeneous areas or ringing artefacts along edges. Alternative solutions based on the use of the total variation can be employed but they often lead to so-called staircase effects. To circumvent these problems and take advantage of both approaches, authors have suggested in [13, 14] to jointly take into account the  $\ell_1$ -norm of the wavelet coefficients and the total variation penalization, which resulted in a significant improvement of the visual restoration quality. These works are however restricted to data corrupted by Gaussian noise.

The purpose of this paper is to extend the work in [13] so as to solve linear inverse problems in the presence of Poisson noise. Let  $\mathcal{H}$  be a real separable Hilbert space corresponding to the frame coefficient space. We aim at finding

$$\min_{x \in C} g(x) + f(x) + h(x) \quad (1)$$

where  $g$ ,  $f$  and  $h$  are assumed to be in the class  $\Gamma_0(\mathcal{H})$  of lower semicontinuous convex functions on  $\mathcal{H}$  taking their values in  $] -\infty, +\infty]$  which are proper (i.e. not identically equal to  $+\infty$ ).  $h$  is related to a total variation penalization and  $C$  is a nonempty closed convex subset of  $\mathcal{H}$ . From a Bayesian interpretation when  $C = \mathcal{H}$ ,  $f + h$  can be viewed as an a priori term on the frame coefficients of the original image,  $g$  corresponds to the Poisson anti log-likelihood (fidelity term) and the main difficulty is to minimize (1) when the degradation is a combined effect of a linear operator (e.g. a blur) and a Poisson noise. These three functions will be detailed in Section 2.

The proposed method is based on recent developments in convex optimization. Its main advantage is that it does not require any approximation of the noise distribution unlike the Anscombe approach [11] or the quadratic extension in [12]. The main contribution of our work consists of taking advantage of the structural properties of the linear degradation operator so as to design an efficient algorithm.

In Section 2, we formulate the minimization problem associated with the restoration of data corrupted by Poisson noise. A hybrid regularization combining a wavelet sparsity promoting term and a total variation term is considered.

The optimization algorithm used to solve this problem is also described based on the work in [13]. In Section 3, we briefly recall the definition of the proximity operator which is the building block of this algorithm. The main difficulty to be addressed here is that the proximity operator of the Kullback-Leibler divergence term does not have an explicit expression. A way of solving this problem is then proposed. Finally, in Section 4, we provide numerical examples as well as algorithm implementation details.

## 2. PROBLEM FORMULATION

The degradation model is the following,

$$z = \mathcal{P}_\alpha(T\bar{y}) \quad (2)$$

where  $\bar{y} \in \mathbb{R}_+^N$  is the original data of size  $N$  degraded by a non-negative valued convolutive operator  $T$  and contaminated by a Poisson noise with scaling parameter  $\alpha \in \mathbb{R}_+^*$ . The vector  $z \in \mathbb{R}^N \subset \mathcal{G} = \mathbb{R}^N$  represents the observed data.

### 2.1 Frame representation

In inverse problems, certain physical properties of the target solution  $\bar{y}$  are most suitably expressed in terms of the coefficients  $\bar{x} = (\bar{\xi}_k)_{k \in \mathbb{K} \subset \mathbb{N}}$  (where  $\mathbb{K} = \{1, \dots, K\}$ ) of its representation  $\bar{y} = \sum_{k \in \mathbb{K}} \bar{\xi}_k e_k$  with respect to a family of vectors  $(e_k)_{k \in \mathbb{K} \subset \mathbb{N}}$  in a Hilbert space  $\mathcal{G}$ . Recall that a family of vectors  $(e_k)_{k \in \mathbb{K}}$  in  $\mathcal{G}$  constitutes a frame if there exist two constants  $\underline{v}$  and  $\bar{v}$  in  $]0, +\infty[$  such that

$$(\forall y \in \mathcal{G}) \quad \underline{v} \|y\|^2 \leq \sum_{k \in \mathbb{K}} |\langle y, e_k \rangle|^2 \leq \bar{v} \|y\|^2. \quad (3)$$

The associated frame operator is the injective bounded linear operator

$$F: \mathcal{G} \rightarrow \ell^2(\mathbb{K}): y \mapsto (\langle y, e_k \rangle)_{k \in \mathbb{K}}, \quad (4)$$

the adjoint of which is the surjective bounded linear operator

$$F^*: \ell^2(\mathbb{K}) \rightarrow \mathcal{G}: (\xi_k)_{k \in \mathbb{K}} \mapsto \sum_{k \in \mathbb{K}} \xi_k e_k. \quad (5)$$

When  $\underline{v} = \bar{v} = v$  in (3),  $(e_k)_{k \in \mathbb{K}}$  is said to be a tight frame. In this case, we have

$$F^* \circ F = v \text{Id}, \quad (6)$$

where Id is the identity operator. A simple example of a tight frame is the union of  $m$  orthonormal bases, in which case  $\underline{v} = \bar{v} = m$ . Considering a frame representation, Model (2) can be re-expressed as,

$$z = \mathcal{P}_\alpha(TF^*\bar{x}) \quad (7)$$

where  $\bar{x}$  represents frame coefficients of the original data ( $\bar{y} = F^*\bar{x}$ ).

### 2.2 Convex optimization

In the context of inverse problems, recent studies proposed to restore the original signal by solving a convex optimization problem of the form:

$$\text{Find } \min_{x \in \mathcal{H}} \sum_{j=1}^J f_j(x) \quad (8)$$

where  $(f_j)_{1 \leq j \leq J}$  are functions of  $\Gamma_0(\mathcal{H})$  where  $\mathcal{H} = \mathbb{R}^K$ .

In restoration methods,  $J$  is usually equal to 2. One function is a smooth term related to the observation model and the second one is a regularization term. In the case of data degraded by a Poisson noise, a standard choice for the first function is a Kullback-Leibler divergence term [15]. More precisely, we employ the Kullback-Leibler divergence  $D_{\text{KL}}$  from  $TF^*x$  to  $z$ , for every  $x \in \mathcal{H}$ :

$$D_{\text{KL}}(TF^*x, z) = \Psi(TF^*x) \quad (9)$$

where,

$$(\forall u = (u^{(n)})_{1 \leq n \leq N} \in \mathcal{G}), \quad \Psi(u) = \sum_{n=1}^N \psi_n(u^{(n)}). \quad (10)$$

and

$$\psi_n(v) = \begin{cases} \alpha v - z^{(n)} + z^{(n)} \ln\left(\frac{z^{(n)}}{\alpha v}\right) & \text{if } z^{(n)} > 0 \text{ and } v > 0, \\ \alpha v & \text{if } z^{(n)} = 0 \text{ and } v \geq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

Concerning the regularization function, in [16, 17, 10] the authors consider a penalization in the wavelet domain corresponding to power functions of the coefficients with exponents  $p_k \in \{1, 4/3, 3/2, 2, 3, 4\}$ . Another type of regularization that can be envisaged employs a total variation measure. Recently, in [13], a hybrid regularization was proposed which gives good results in the context of an additive white Gaussian noise. We thus propose to use a similar compound regularization in the case when the data are corrupted by Poisson noise. In this context, Problem (1) is solved by setting  $g = \Psi \circ T \circ F^*$  where  $\Psi$  is defined by (10). The function  $f$  corresponds to the regularization term operating in the wavelet domain, which is chosen such that  $\forall x = (x^{(k)})_{1 \leq k \leq K} \in \mathcal{H}$ ,  $f(x) = \sum_{k=1}^K \phi_k(x^{(k)})$  where, for every  $k \in \{1, \dots, K\}$ ,  $\phi_k$  is a finite function of  $\Gamma_0(\mathbb{R})$  such that  $\lim_{|x^{(k)}| \rightarrow +\infty} \phi_k(x^{(k)}) = +\infty$  and  $h$  represents a total variation term such that  $h = \text{tv} \circ F^*$ , tv being the total variation operator defined in [13, Experiment 2]. Hence, the considered minimization problem in (1) becomes,

$$\min_{x \in \mathcal{H}} D_{\text{KL}}(TF^*x, z) + \vartheta f(x) + \kappa \text{tv}(F^*x) + \iota_C(x) \quad (11)$$

where  $\iota_C$  is the indicator function of a closed convex set  $C$  (for example related to support or value range constraints) such that,

$$(\forall x \in \mathcal{H}) \quad \iota_C(x) = \begin{cases} 0, & \text{if } x \in C; \\ +\infty, & \text{otherwise.} \end{cases} \quad (12)$$

Throughout the paper, it is assumed that  $(TC^*) \cap ]0, +\infty[^N \neq \emptyset$  with  $C^* = F^*C = \{F^*x \mid x \in C\}$ .

The nonnegative real parameters  $\vartheta$  and  $\kappa$  control the degree of smoothness in the wavelet and in the space domain, respectively. Notice that, when  $\vartheta = 0$ , the regularization reduces to the standard total variation penalization and that the approach we propose in this paper thus provides also an efficient numerical solution to such a more classical penalized optimization problem.

In the class of convex optimization methods, to the best of our knowledge, only one algorithm allows us to efficiently

minimize the sum of four terms, which are not differentiable. The algorithm was proposed by Combettes and Pesquet in [13] and is summarized next.

**Algorithm 2.1** Let  $\gamma \in ]0, +\infty[$ . For every  $j \in \{1, \dots, J\}$ , set  $(\omega_j)_{1 \leq j \leq J} \in ]0, 1]^J$  such that  $\sum_{j=1}^J \omega_j = 1$ ,  $(y_{j,0})_{1 \leq j \leq J} \in \mathcal{H}^J$  and  $x_0 = \sum_{j=1}^J \omega_j y_{j,0}$ . Let  $(a_{j,l})_{l \in \mathbb{N}}$  be a sequence in  $\mathcal{H}$  which corresponds to possible errors in the computation of proximity operators. Then, the sequence  $(x_l)_{l \geq 1}$  is generated by the following routine: for every  $l \in \mathbb{N}$ ,

$$\left\{ \begin{array}{l} \text{For } j = 1, \dots, J \\ \quad | p_{j,l} = \text{prox}_{\gamma/\omega_j f_j} y_{j,l} + a_{j,l} \\ p_l = \sum_{j=1}^J \omega_j p_{j,l} \\ \lambda_l \in ]0, 2[ \\ \text{For } j = 1, \dots, J \\ \quad | y_{j,l+1} = y_{j,l} + \lambda_l (2 p_l - x_l - p_{j,l}) \\ x_{l+1} = x_l + \lambda_l (p_l - x_l) \end{array} \right. \quad (13)$$

The prox operator introduced in the first loop will be defined Section 3. The sequence  $(x_l)_{l \geq 1}$  converges to a solution to Problem (8) under the following assumption.

**Assumption 2.1**

- (i)  $\lim_{\|x\| \rightarrow +\infty} f_1(x) + \dots + f_J(x) = +\infty$ .
- (ii)  $\bigcap_{j=1}^J \text{rint dom } f_j \neq \emptyset$  (when  $\mathcal{H}$  is finite-dimensional)<sup>1</sup>.
- (iii)  $\sum_{l \in \mathbb{N}} \lambda_l (2 - \lambda_l) = +\infty$ .
- (iv)  $(\forall j \in \{1, \dots, J\}) \quad \sum_{l \in \mathbb{N}} \lambda_l \|a_{j,l}\| < +\infty$ .

The main difficulty in applying this algorithm to our restoration problem is that it requires to compute the proximity operators associated to each of the four terms in (11). Closed forms of these quantities are known for the indicator function, the function  $f$  [10] and the total variation term (through the decomposition method proposed in [13]). The main problem remains in the computation of the proximity operator of the Kullback-Leibler divergence term. In the next section, we will recall the notion of proximity operator, before providing an answer to this question.

### 3. CONVEX OPTIMIZATION TOOLS: PROXIMITY OPERATOR

#### 3.1 Definition and properties

A fundamental tool in the study of the convex optimization methods is the proximity operator introduced by Moreau in 1962 [18]. The proximity operator of  $\varphi \in \Gamma_0(\mathcal{H})$  is defined by

$$\text{prox}_\varphi : \mathcal{H} \rightarrow \mathcal{H}; x \mapsto \arg \min_{y \in \mathcal{H}} \frac{1}{2} \|y - x\|^2 + \varphi(y). \quad (14)$$

We thus see that  $\text{prox}_{\text{ic}}$  reduces to the projection  $P_C$  onto the convex set  $C$ . Other examples of proximity operators corresponding to the potential functions of standard log-concave univariate probability densities have been listed in [10].

We now recall the proximity operator of the potential associated with a Gamma distribution which is closely related to the Kullback-Leibler divergence [19].

<sup>1</sup>The relative interior of a set  $S$  of  $\mathcal{H}$  is designated by  $\text{rint } S$  and the domain of a function  $f : \mathcal{H} \rightarrow ]-\infty, +\infty]$  is  $\text{dom } f = \{x \in \mathcal{H} | f(x) < +\infty\}$ .

**Example 3.1** [10] Let  $\omega \in ]0, +\infty[$ ,  $\chi \in [0, +\infty[$ , and set

$$\varphi : \mathbb{R} \rightarrow ]-\infty, +\infty]: \eta \mapsto \begin{cases} -\chi \ln(\eta) + \omega \eta, & \text{if } \eta > 0; \\ +\infty, & \text{if } \eta \leq 0. \end{cases} \quad (15)$$

Then, for every  $\eta \in \mathbb{R}$ ,

$$\text{prox}_\varphi \eta = \frac{\eta - \omega + \sqrt{|\eta - \omega|^2 + 4\chi}}{2}. \quad (16)$$

Our minimization problem being formulated in a frame representation, we will also need a property concerning the calculation of the proximity operator of the composition of a function of  $\Gamma_0(\mathcal{G})$  and a linear operator:

**Proposition 3.2** [19] Let  $\mathcal{H}$  and  $\mathcal{G}$  be real Hilbert spaces, let  $\varphi \in \Gamma_0(\mathcal{G})$ , and let  $L : \mathcal{H} \rightarrow \mathcal{G}$  be a bounded linear operator. Suppose that the composition of  $L$  and  $L^*$  satisfies  $L \circ L^* = \zeta \text{Id}$ , for some  $\zeta \in ]0, +\infty[$ . Then  $\varphi \circ L \in \Gamma_0(\mathcal{H})$  and

$$\text{prox}_{\varphi \circ L} = \text{Id} + \zeta^{-1} L^* \circ (\text{prox}_{\zeta \varphi} - \text{Id}) \circ L. \quad (17)$$

#### 3.2 Case of a convolutive operator $T$

In order to solve Problem (11), we will be interested in determining the proximity operator of  $g = \Psi \circ T \circ F^*$ . As will be shown next, the proximity operator of this function can be determined in a closed form for specific cases only. However,  $g$  can be decomposed as a sum of functions for which the proximity operators can be calculated explicitly. We subsequently assume that:

**Assumption 3.3**  $(e_k)_{1 \leq k \leq K}$  is a tight frame of  $\mathcal{G} = \mathbb{R}^N$  with frame constant  $\nu > 0$ .

Let us now focus on function  $g$ . Let  $(\mathbb{I}_i)_{1 \leq i \leq I}$  be a partition of  $\{1, \dots, N\}$  in nonempty sets. For every  $i \in \{1, \dots, I\}$ , let  $M_i$  be the number of elements in  $\mathbb{I}_i$  and let  $\Upsilon_i : \mathbb{R}^{M_i} \rightarrow ]0, +\infty[ : (\eta_n)_{n \in \mathbb{I}_i} \mapsto \sum_{n \in \mathbb{I}_i} \psi_n(\eta_n)$ . We have then  $g = \sum_{i=1}^I \Upsilon_i \circ T_i \circ F^*$  where  $T_i$  is the linear operator from  $\mathbb{R}^N$  to  $\mathbb{R}^{M_i}$  associated with the matrix

$$[t_{m_1} \dots t_{m_{M_i}}]^\top$$

with  $\mathbb{I}_i = \{m_1, \dots, m_{M_i}\}$ . The following assumption will play a prominent role in the rest of paper:

**Assumption 3.4** For all  $i \in \{1, \dots, I\}$ ,  $(t_n)_{n \in \mathbb{I}_i}$  is a family of orthogonal vectors having the same norm  $\sigma_i^{1/2}$  with  $\sigma_i > 0$ .

Note that this assumption is obviously satisfied when  $I = N$ , that is when  $\forall i \in \{1, \dots, I\}$ ,  $\mathbb{I}_i$  reduces to a singleton.

**Proposition 3.5** Under Assumptions 3.3 and 3.4, we have,  $(\forall i \in \{1, \dots, I\}) (\forall x \in \mathcal{H})$

$$\text{prox}_{\Upsilon_i \circ T_i \circ F^*}(x) = x + \frac{1}{\nu \sigma_i} F T_i^* (\pi_i^{(n)} - \eta_i^{(n)})_{n \in \mathbb{I}_i} \quad (18)$$

where  $(\eta_i^{(n)})_{n \in \mathbb{I}_i} = T_i F^* x$  and

$$(\forall n \in \mathbb{I}_i) \quad \pi_i^{(n)} = \frac{\eta_i^{(n)} - \alpha \nu \sigma_i + \sqrt{|\eta_i^{(n)} - \alpha \nu \sigma_i|^2 + 4 \nu \sigma_i z^{(n)}}}{2}. \quad (19)$$

**Remark**

- (i) This result is a consequence of Proposition 3.2 by setting  $L = T_i F^*$  and by using Example 3.1 to derive  $\text{prox}_{v\sigma_i \Upsilon_i}$ .
- (ii) It can be noticed that the application of  $T_i$  or  $T_i^*$  reduces to standard operations in signal processing. The application of  $T_i$  consists of two steps: a convolution with the degradation filter and a decimation for selected locations ( $n \in \mathbb{I}_i$ ). The application of  $T_i^*$  also consists of two steps: an interpolation step (by setting zero everywhere except for indices  $n \in \mathbb{I}_i$ ) followed by a convolution with the filter with conjugate frequency response.

To reduce the number of proximity operators to be computed, we want to find the smallest integer  $I$  such that, for every  $i \in \{1, \dots, I\}$ ,  $(t_n)_{n \in \mathbb{I}_i}$  is an orthogonal basis. Consider the particular case of a deconvolution problem for images of size  $N_1 \times N_2$  ( $N = N_1 N_2$ ) involving a periodic convolution where  $Q_1 \times Q_2$  is the kernel size. The operator  $T$  is associated with a matrix having a circulant-block circulant structure [20] and consequently,

$$\begin{bmatrix} T_1 \\ \vdots \\ T_N \end{bmatrix} = \begin{bmatrix} \Theta_0 & O & \dots & O & \Theta_{Q_1-1} & \dots & \Theta_1 \\ \Theta_1 & \Theta_0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \Theta_{Q_1-1} \\ \Theta_{Q_1-1} & \ddots & \ddots & \ddots & \ddots & \ddots & O \\ O & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & O \\ O & \dots & O & \Theta_{Q_1-1} & \dots & \Theta_1 & \Theta_0 \end{bmatrix} \quad (20)$$

where  $O$  is the null matrix of size  $N_2 \times N_2$  and ( $\forall q_1 \in \{0, \dots, Q_1 - 1\}$ )  $\Theta_{q_1}$  is the matrix of size  $N_2 \times N_2$  defined by

$$\Theta_{q_1} = \begin{bmatrix} \theta_{q_1,0} & 0 & \dots & 0 & \theta_{q_1,Q_2-1} & \dots & \theta_{q_1,1} \\ \theta_{q_1,1} & \theta_{q_1,0} & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \theta_{q_1,Q_2-1} \\ \theta_{q_1,Q_2-1} & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \theta_{q_1,Q_2-1} & \dots & \theta_{q_1,1} & \theta_{q_1,0} \end{bmatrix} \quad (21)$$

$(\theta_{q_1,q_2})_{0 \leq q_1 < Q_1, 0 \leq q_2 < Q_2}$  denotes here the point spread function of the degradation filter.

Let us define  $I_1 = \min\{i_1 \geq Q_1 \mid N_1 = 0 \text{ mod } i_1\}$  and,  $I_2 = \min\{i_2 \geq Q_2 \mid N_2 = 0 \text{ mod } i_2\}$ . In order to satisfy Assumption 3.4, we subsequently set  $I = I_1 I_2$ .

For all  $(i_1, i_2) \in \{1, \dots, I_1\} \times \{1, \dots, I_2\}$ , set  $i = i_2 + I_2(i_1 - 1)$  and define

$$\mathbb{I}_i = \{n_2 + N_2(n_1 - 1) \mid (n_1, n_2) \in \{1, \dots, N_1\} \times \{1, \dots, N_2\} \text{ and } n_1 = i_1 \text{ mod } I_1, n_2 = i_2 \text{ mod } I_2\}. \quad (22)$$

Then, Assumption 3.4 holds and, for all  $i \in \{1, \dots, I\}$ ,  $\sigma_i = \sum_{q_1=0}^{Q_1-1} \sum_{q_2=0}^{Q_2-1} |\theta_{q_1,q_2}|^2$ .

We have therefore defined sets  $(\mathbb{I}_i)_{1 \leq i \leq I}$  for which we can compute the associated proximity operators as expressed by

Proposition 3.5. Note that calculations for other forms of convolution operations (e.g. including zero-padding) can be achieved in a similar way.

Regarding Algorithm 2.1, the  $J$  proximity operators to be computed are: the  $I$  proximity operators corresponding to  $\text{prox}_{\frac{\gamma}{\omega_i} \Upsilon_i \circ T_i \circ F^*}$ , the four proximity operators associated with the total variation term [13], the proximity operator of the regularization function  $f$  and also, the projection onto the convex set  $C$ . To sum up,  $J = I + 6$  proximity operator computations are needed in order to solve Problem (11).

**4. EXPERIMENTAL RESULTS**

In our simulations, our objectives are twofold: we will first be interested in studying the influence of the combination of total variation and wavelet regularization terms and then, we will compare the results obtained by using the proposed algorithm with those corresponding to state-of-the-art methods. Two test images ( $N_1 = N_2 = 256$ ) will be considered (see Fig. 1). For both examples,  $T$  is a uniform blur with kernel dimensions  $Q_1 = Q_2 = 3$ . Therefore, the partition cardinality  $I$  introduced in Section 3.2 is such that  $I_1 = I_2 = 4$ .  $C$  is here defined as  $(F^*)^{-1} C^*$  with  $C^* = [0, 255]^N$ . A tight frame version of the dual-tree transform (DTT) proposed in [21] ( $v = 2$ ) using Symlets of length 6 has been employed over 3 resolution levels. Strictly convex non-differentiable potential functions are chosen, such that  $\phi_k = \omega_k |\cdot|^{p_k} + \chi_k |\cdot|$  where  $(\omega_k, \chi_k) \in ]0, +\infty[^2$  and  $p_k \in \{4/3, 3/2, 2\}$ .

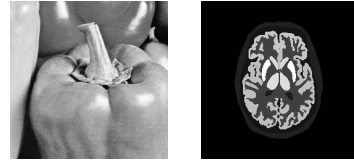


Figure 1: “Peppers” image (left) and “Sebal” image (right).

**4.1 Influence of each regularization term**

We present some numerical and visual results considering different adjustments of  $\vartheta$  and  $\kappa$ . This experiment allows us to illustrate the influence of the wavelet regularization and the total variation one. In the images displayed in Figure 2, it can be observed the artefacts related to the wavelet regularization, the staircase effects which are typical of the total variation penalization and also the advantage of using a hybrid regularization. Table 1 provides quantitative results allowing us to evaluate the impact of the adjustment of the regularization factors  $\vartheta$  and  $\kappa$ .

	$\kappa = 0.04$	$\kappa = 0.02$	$\kappa = 0.01$	$\kappa = 0.005$
$\vartheta = 1$	20.6	21.6	21.9	21.9
$\vartheta = 0.2$	21.4	22.6	22.9	22.5
$\vartheta = 0.1$	21.6	22.8	<b>23.1</b>	22.1
$\vartheta = 0.05$	21.8	23.0	23.0	21.4
$\vartheta = 0.01$	21.8	22.8	22.1	19.1

Table 1: SNR for “Peppers” image with  $\alpha = 0.1$ . Iteration number (IN) lies between 100 and 200.

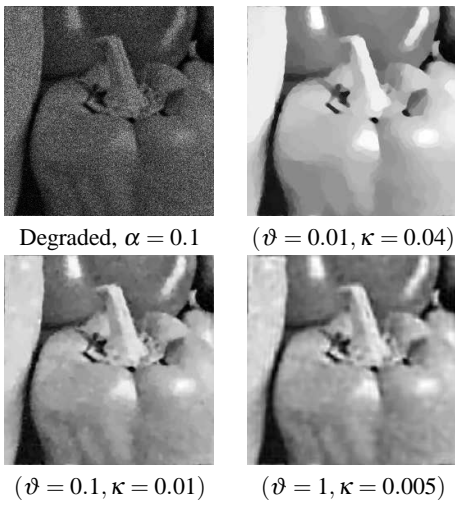


Figure 2: Some results for “Peppers” image.

## 4.2 Comparison with existing methods

We now aim at comparing the proposed algorithm with existing methods such as the regularized Expectation-Maximization proposed by Byrne [9], the Anscombe approximation method proposed by Dupé *et al.* [11] and the quadratic extension proposed in [12] (for optimal parameter values). The results are given in Table 2 for “Sebal” image and for different noise intensity factor  $\alpha$ . Whatever the scaling parameter chosen, the proposed approach always gives better results in terms of Signal Noise Ratio (SNR). In addition, the improvement in visual quality obtained by adding the total variation penalization to the wavelet regularization is illustrated in Figure 3.

$\alpha$	0.01	0.05	0.1	0.5
EM-Reg. [9]	4.12	6.87	8.12	11.1
Anscombe - DTT [11]	7.93	10.8	11.7	13.6
Quadratic ext. - DTT [12]	8.91	11.0	11.8	13.6
Proposed algorithm - DTT	<b>10.1</b>	<b>12.0</b>	<b>13.3</b>	<b>15.6</b>

Table 2: SNR for “Sebal” image.

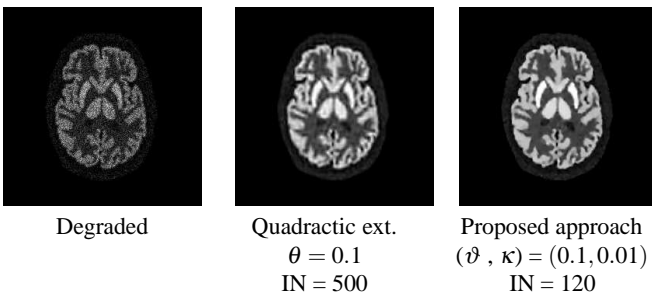


Figure 3: Some visual results for “Sebal” image with  $\alpha = 0.1$ .

## 5. CONCLUSION

A new approach to restore data degraded by a convolution and Poisson noise has been proposed. The main advan-

tages of the method are (i) to deal directly with Kullback-Leibler divergence (without requiring any approximation); (ii) to permit the use of sophisticated regularization terms, e.g. one promoting sparsity in a wavelet frame and a total variation penalization. Numerical and visual results demonstrate the effectiveness of the proposed approach.

## REFERENCES

- [1] F. J. Anscombe. The transformation, of Poisson, binomial and negative-binomial data. *Biometrika*, 35(3/4):246–254, Dec. 1948.
- [2] P. Fryzlewicz and G. P. Nason. A Haar-Fisz algorithm for poisson intensity estimation. *Jour. Comput. and Graph. Stat.*, 13:621–638, 2004.
- [3] C. A. Bouman and K. Sauer. A unified approach to statistical tomography using coordinate descent optimization. *IEEE Trans. on Image Proc.*, 5(3):480–492, Mars 1996.
- [4] J. Verhaeghe, D. Van De Ville, I. Khalidov, Y. D’Asseler, I. Lemahieu, and M. Unser. Dynamic PET reconstruction using wavelet regularization with adapted basis functions. *IEEE Trans. on Medical Imaging*, 27(7):943–959, July 2008.
- [5] W. H. Richardson. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America*, 62(1):55–59, 1972.
- [6] L. B. Lucy. An iterative technique for the rectification of observed distributions. *Astronomical Journal*, 79(6):745–754, 1974.
- [7] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1):259–268, 1992.
- [8] N. Dey, L. Blanc-Féraud, C. Zimmer, Z. Kam, P. Roux, J.C. Olivo-Marin, and J. Zerubia. Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution. *Microscopy Research Technique*, 69:260–266, 2006.
- [9] C. L. Byrne. Iterative image reconstruction algorithms based on cross-entropy minimization. *IEEE Trans. on Image Proc.*, 2(1):96–103, Jan. 1993.
- [10] C. Chaux, P. L. Combettes, J.-C. Pesquet, and V. R. Wajs. A variational formulation for frame-based inverse problems. *Inverse Problems*, 23:1495–1518, June 2007.
- [11] F.-X. Dupé, M. J. Fadili, and J.-L. Starck. Image deconvolution under Poisson noise using sparse representations and proximal thresholding iteration. In *Proc. Int. Conf. on Acoust., Speech and Sig. Proc.*, pages 761–764, Las Vegas, Mar. 30-Apr. 4, 2008.
- [12] C. Chaux, J.-C. Pesquet, and N. Pustelnik. Nested iterative algorithms for convex constrained image recovery problems. *SIAM Journal on Imaging Sciences*, 2009. To appear.
- [13] P. L. Combettes and J.-C. Pesquet. A proximal decomposition method for solving convex variational inverse problems. *Inverse Problems*, 24(6), Dec. 2008.
- [14] J. M. Bioucas-Dias and M. A. T. Figueiredo. An iterative algorithm for linear inverse problems with compound regularizers. *Proc. Int. Conf. on Image Processing*, 12–15 Oct. 2008. 4 p.
- [15] D. Titterton. On the iterative image space reconstruction algorithm for ECT. *IEEE Trans. Med. Imaging*, MI-6:52–56, 1987.
- [16] M. A. T. Figueiredo and R. D. Nowak. An EM algorithm for wavelet-based image restoration. *IEEE Trans. on Image Proc.*, 12(8):906–916, August 2003.
- [17] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Applied Math.*, 57:1413–1457, 2004.
- [18] J. J. Moreau. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France*, 93:273–299, 1965.
- [19] P. L. Combettes and J.-C. Pesquet. A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE Journal of Selected Topics in Sig. Proc.*, 1(4):564–574, Dec. 2007.
- [20] G. H. Golub and C. F. Van Loan. *Matrix computations*. The Johns Hopkins University Press; 3rd edition, 1996.
- [21] C. Chaux, L. Duval, and J.-C. Pesquet. Image analysis using a dual-tree  $M$ -band wavelet transform. *IEEE Trans. on Image Proc.*, 15(8):2397–2412, Aug. 2006.