A WAVELET-BASED QUADRATIC EXTENSION METHOD FOR IMAGE DECONVOLUTION 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

Iterative optimization algorithms such as the forward-backward and Douglas-Rachford algorithms have recently gained much popularity since they provide efficient solutions to a wide class of non-smooth convex minimization problems arising in signal/image recovery. However, when images are degraded by a convolution operator and a Poisson noise, a particular attention must be paid to the associated minimization problem. To solve it, we propose a new optimization method which consists of two nested iterative steps. The effectiveness of the proposed method is demonstrated via numerical comparisons.

***Index Terms*—** Deconvolution, iterative methods, Poisson distributions, optimization methods, wavelet transforms.

1. INTRODUCTION

In many application areas, image restoration plays a primary role. When data are degraded by a convolutive blur and the addition of Gaussian noise, a large panel of methods can be used for restoration purposes. However, in real applications, the nature of the encountered noise often differs from a Gaussian one. For example, Poisson noise (which is dependent on image intensity) is found in tomography, astronomy and microscopy. This kind of noise may appear much more challenging to cope with. In this work, we consider the following degradation model:

$$z = \mathcal{P}_\alpha(T\overline{y}) \tag{1}$$

where $\overline{y}$ is the original image degraded by a blur operator T and contaminated by a Poisson noise with scaling factor α. The effect of the noise is denoted by $\mathcal{P}_\alpha$.

The simplest way to take into account Poisson noise is to process the data as if they were corrupted with Gaussian noise and consequently, to apply standard restoration tools [1, 2]. A better approach consists of applying a pre-processing (also called Variance Stabilizing Transform) on the data. For example, the Anscombe transform [3] can be used to pre-process Poisson data and then, to restore the transformed data as if they were corrupted by Gaussian noise. Fryzlewicz and Nason explored another stabilizing variance approach in [4] named the Haar-Fisz transform. The main drawback of this pre-processing lies in the fact that it uses the Haar decomposition which is known to be especially efficient for blockwise constant images. More recently, iterative algorithms based on a variational approach were proposed in the Poisson denoising context. In [5], Sardy et al. developed an ℓ_1-penalized likelihood approach and more recently, in [6], the authors used the Douglas-Rachford algorithm within a convex optimization framework. For deconvolution problems, algorithms based on the minimization of the Kullback-Leibler divergence were studied in [7]. An extension of this approach was investigated in [8] to incorporate a total variation penality term. More recently, a block iterative method for binary tomographic reconstruction based on convex contraint sets was proposed in [9]. In [10, 11], the authors considered algorithms mixing forward-backward [12] and Douglas-Rachford steps. On the one hand, Dupé et al. [10] investigated an elegant adaptation of the Anscombe approach. On the other hand, a quadratic extension dealing with signal dependent Gaussian noise was introduced in [11]. Note that a different polynomial approximation of the objective function was proposed in [13]. The objective of this paper is to adapt the extension proposed in [11] in order to deal with Poisson distributions and, to investigate its properties. Notice that the proposed method relies on a wavelet-like representation as in some of the aforementioned approaches.

The paper is organized as follows: in Section 2, we formulate the restoration problem from a frame representation perspective. A Bayesian interpretation of the related convex minimization problem is given. Then, we propose a quadratic extension technique to circumvent the problem of the non Lipschitz differentiability of the likelihood of the Poisson data. In Section 3, we present the new algorithm combining a forward-backward and a Dykstra's step so as to restore blurred images contaminated by Poisson noise. Finally, the effectiveness of the proposed approach is demonstrated via simulation results.

2. IMAGE DECONVOLUTION IN THE PRESENCE OF POISSON NOISE

2.1. Problem statement

We consider the degradation model (1). An efficient probabilistic prior model on the unknown image $\overline{y}$ is adopted by considering a frame representation of this image [12]. We thus use a linear representation of the form $\overline{y} = F^*\overline{x}$ where $F^* : \mathcal{H} \to \mathcal{G}$ is a tight frame synthesis operator with $\mathcal{H} = \mathbb{R}^K$, $\mathcal{G} = \mathbb{R}^N$ and $K \geq N$. This means that $F^* \circ F = \nu \mathrm{Id}$ with $\nu \in]0, +\infty[$. In this context, (1) can be reformulated as $z = \mathcal{P}_\alpha(TF^*\overline{x})$ where $\overline{x}$ is the unknown vector of frame coefficients.

2.2. Bayesian framework

The image $\overline{u} = T\overline{y}$ (resp. z) is viewed as a realization of a nonnegative real-valued random vector $\overline{U} = (\overline{U}^{(i)})_{1\leq i\leq N}$ (resp. $Z = (Z^{(i)})_{1\leq i\leq N}$). Conditionally to $\overline{U} = (u^{(i)})_{1\leq i\leq N} \in \mathcal{G}$, the random vector Z is assumed to have independent components with conditional probability: $\forall z^{(i)} \in \mathbb{N}$,

$$\mathsf{P}(Z^{(i)} = z^{(i)} \mid \overline{U}^{(i)} = u^{(i)}) = \frac{(\alpha u^{(i)})^{z^{(i)}}}{z^{(i)}!} \exp(-\alpha u^{(i)}) \quad (2)$$

where $\alpha \in]0, +\infty[$ is a scaling parameter.

We then assume that the vector $\overline{x}$ of frame coefficients is a realization of a random vector $\overline{X}$ with independent components and, each component $\big(\overline{X}^{(k)}\big)_{1\leq k\leq K}$ has a probability density given by: $\forall x^{(k)} \in \mathbb{R}$, $p_{\overline{X}^{(k)}}(x^{(k)}) \propto \exp(-\phi_k(x^{(k)}))$ where ϕ_k is a finite convex function allowing us to model frame coefficients. A large choice of potential functions $(\phi_k)_{1\leq k\leq K}$ is possible as shown in [12]. Under these assumptions, it is readily shown that a MAP estimator of the vector of frame coefficients $\overline{x}$ can be obtained from z by solving:

$$\min_{x=(x^{(k)})_{1\leq k\leq K}\in\mathcal{H}} \sum_{k=1}^{K} \phi_k(x^{(k)}) - \ln \mathsf{P}(Z = z \mid \overline{U} = TF^*x).$$

2.3. Link with convex optimization

One generic problem in convex optimization is to determine:

$$\min_{x\in C} \quad f(x) + g(x) \quad (3)$$

where C is a closed convex subset of $\mathcal{H}$, f and g are functions in $\Gamma_0(\mathcal{H})$. $\Gamma_0(\mathcal{H})$ denotes the class of lower semicontinuous convex functions taking their values in $]-\infty, +\infty]$.

In our restoration problem, g denotes the fidelity term related to the Poisson distribution of the noise and f the a priori term on $\overline{x}$. Thus, g is chosen as, $\forall x \in \mathcal{H}$, $g(x) = \Psi(TF^*x)$, where

$$\big(\forall u = \big(u^{(i)}\big)_{1\leq i\leq N} \in \mathcal{G}\big) \qquad \Psi(u) = \sum_{i=1}^{N} \psi_i\big(u^{(i)}\big). \quad (4)$$

From (2), it can be deduced that: $\forall \upsilon \in \mathbb{R}$,

$$\psi_i(\upsilon) = \begin{cases} \alpha\upsilon - z^{(i)} + z^{(i)} \ln\left(\dfrac{z^{(i)}}{\alpha\upsilon}\right) & \text{if } z^{(i)} > 0 \text{ and } \upsilon > 0, \\ \alpha\upsilon & \text{if } z^{(i)} = 0 \text{ and } \upsilon \geq 0, \\ +\infty & \text{otherwise.} \end{cases} \quad (5)$$

Furthermore, the prior f is defined as:

$$(\forall x = \big(x^{(k)}\big)_{1\leq k\leq K} \in \mathbb{R}^K) \quad f(x) = \sum_{k=1}^{K} \phi_k\big(x^{(k)}\big). \quad (6)$$

Finally, as it is customary in image processing, a constraint on the range of the pixel values is introduced by setting

$$C = \big\{x \in \mathcal{H} \;\big|\; F^*x \in [0, 255]^N\big\}. \quad (7)$$

2.4. Lipschitz differentiability issue

The functions f and g as defined above are in $\Gamma_0(\mathcal{H})$ but an additional assumption is often required to ensure the convergence of optimization algorithms such as the modified forward-backward algorithm in [11], namely g should be β-Lipschitz differentiable on C with $\beta \in]0, +\infty[$. Unfortunately the assumption is not satisfied by g.

In [11], a quadratic extension was proposed so as to improve convergence profiles. In the present case, we will show that this technique can be adapted in order to obtain a lower approximation g_θ of g, which is Lipschitz differentiable on C. This function is defined as: $g_\theta = \Psi_\theta \circ T \circ F^*$, where $\theta \in]0, +\infty[$ and Ψ_θ is given by

$$\big(\forall u = \big(u^{(i)}\big)_{1\leq i\leq N} \in \mathcal{G}\big) \qquad \Psi_\theta(u) = \sum_{i=1}^{N} \psi_{\theta,i}\big(u^{(i)}\big). \quad (8)$$

The functions $(\psi_{\theta,i})_{1\leq i\leq N}$ are here defined as: $\forall \upsilon \in \mathbb{R}$, $\psi_{\theta,i}(\upsilon) =$

$$\begin{cases} \dfrac{\theta}{2}\upsilon^2 + \zeta_{i,1}(\theta)\,\upsilon + \zeta_{i,0}(\theta) & \text{if } z^{(i)} > 0 \text{ and } -\epsilon(\theta) \leq \upsilon < \upsilon_i(\theta) \\ \alpha\upsilon & \text{if } z^{(i)} = 0 \text{ and } -\epsilon(\theta) \leq \upsilon < 0 \\ \psi_i(\upsilon) & \text{otherwise,} \end{cases}$$

where $\upsilon_i(\theta) = (z^{(i)}/\theta)^{1/2}$ is such that $0 \leq \psi_i''(\upsilon) \leq \theta \Leftrightarrow \upsilon \geq \upsilon_i(\theta)$, and the constants $\zeta_{i,0}(\theta)$ and $\zeta_{i,1}(\theta)$ are chosen so as to ensure the continuity of $\psi_{\theta,i}$ in $\upsilon_i(\theta)$. The validity of this approximation is secured by the following result:

Proposition 2.1 [14] *Assume that*

(i) *ϵ is a decreasing positive function such that* $\lim_{\theta\to+\infty} \epsilon(\theta) = 0$,

(ii) $T[0, 255]^N \subset [0, +\infty[^N$,

(iii) *f is coercive or C is bounded,*

(iv) *f is strictly convex on C.*

Then, there exists $\overline{\theta} \in]0, +\infty[$ such that, for every $\theta \in [\overline{\theta}, +\infty[$, $f + g_\theta$ has a unique minimizer on the convex set C which is the minimizer of $f + g$ on C.

The proposed quadratic extension is illustrated in Fig. 1. When θ becomes greater than some lower value $\overline{\theta}$, the solution of the original minimization problem is perfectly found as g is approximated by g_θ. The choice of θ will be discussed in Section 4.

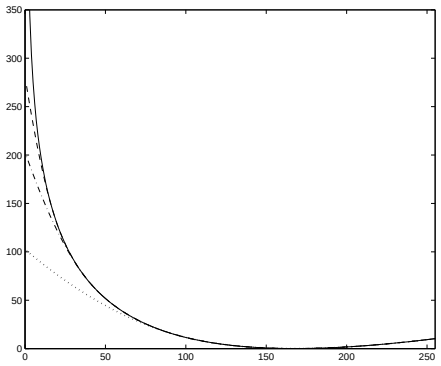


Fig. 1. Quadratic approximation: Poisson likelihood (continuous line) and its extensions for $\theta = 0.3$ (dashed line), $\theta = 0.07$ (dash-dot line) and $\theta = 0.01$ (dotted line).

3. PROPOSED ALGORITHM

3.1. Optimization background

The approximated version of Problem (3) can be expressed as:

$$\min_{x \in \mathcal{H}} f(x) + g_\theta(x) + \iota_C(x) \tag{9}$$

where ι_C denotes the indicator function of C, *i.e.*, $\forall x \in \mathcal{H}$, $\iota_C(x) = 0$ if $x \in C$, $+\infty$ otherwise. Among convex optimization methods, the forward-backward algorithm appears as an appropriate solution for minimizing (9) when f is non necessarily smooth. It is based on a useful tool called the proximity operator. We recall that the proximity operator of $f \in \Gamma_0(\mathcal{H})$ is $\mathrm{prox}_f \colon \mathcal{H} \to \mathcal{H} \colon x \mapsto \mathrm{prox}_f x$ where $\mathrm{prox}_f x$ is the unique minimizer of $f + \frac{1}{2}\|. - x\|^2$. This operator generalizes the notion of projection as $\mathrm{prox}_{\iota_C} = P_C$ where P_C denotes the projection onto C.

3.2. Forward-backward approach

Let $x_0 \in C$ be an initial image. The algorithm constructs a sequence $(x_n)_{n \geq 1}$ by the iteration: for every $n \in \mathbb{N}$,

$$x_{n+1} = x_n + \lambda_n \big(\mathrm{prox}_{\iota_C + \gamma_n f}(x_n - \gamma_n \nabla g_\theta(x_n)) - x_n \big)$$

where $\gamma_n \in]0, 2(\nu\theta\|T\|^2)^{-1}[$ is the algorithm step-size and $\lambda_n \in]0, 1[$ is a relaxation parameter. Under these conditions, it is known that, if $(\forall n \in \mathbb{N})\ x_n \in C$, then $(x_n)_{n \in \mathbb{N}}$ converges to a solution to Problem (9) [12, Theorem 5.4].

3.3. Dykstra's algorithm

A main difficulty in the application of the forward-backward algorithm is the determination of $\mathrm{prox}_{\iota_C + \gamma_n f}$ [11]. To perform this task, we propose to use Dykstra's algorithm [15]. Set $r_0 \in \mathcal{H}$ and $p_0 = q_0 = 0$. Sequences $(r_m)_{m \geq 1}$ and $(s_m)_{m \geq 1}$ are generated by the following routine: $\forall m \in \mathbb{N}$,

$$(s_m, p_{m+1}, r_{m+1}, q_{m+1}) = \mathcal{D}_f(r_m, p_m, q_m)$$
$$\Leftrightarrow \begin{cases} s_m & = P_C(r_m + p_m) \\ p_{m+1} & = r_m + p_m - s_m \\ r_{m+1} & = \mathrm{prox}_f(s_m + q_m) \\ q_{m+1} & = s_m + q_m - r_{m+1}. \end{cases}$$

Then $(r_m)_{m \in \mathbb{N}}$ and $(s_m)_{m \in \mathbb{N}}$ converge to the solution of $\mathrm{prox}_{f + \iota_C}(r_0)$ [15, Theorem 3.3].

3.4. Resulting algorithm

The proposed method consists of the following combination of the two previous algorithms:

① Set $x_0 \in C$ and $n = 0$.
② Set $r_{n,0} = x_n - \gamma_n \nabla g_\theta(x_n)$ and $p_{n,0} = q_{n,0} = 0$.
③ For $m = 0, \ldots, M_n$
  a) $(s_{n,m}, p_{n,m+1}, r_{n,m+1}, q_{n,m+1})$
  $= \mathcal{D}_{\gamma_n f}(r_{n,m}, p_{n,m}, q_{n,m})$
  b) If $s_{n,m} = s_{n,m-1}$, goto ④.
④ Set $x_{n+1} = x_n + \lambda_n \big(s_{n,m} - x_n \big)$.
⑤ Increment n ($n \leftarrow n + 1$) and goto ②.

The following result can then be deduced from the convergence properties of the forward-backward and Dykstra's algorithms.

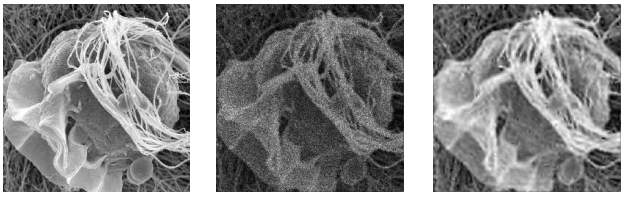
Proposition 3.1 *Under Assumptions (ii) and (iii) of Proposition 2.1, there exists a sequence of positive integers $(\overline{M}_n)_{n \in \mathbb{N}}$ such that, if $(\forall n \in \mathbb{N})\ M_n \geq \overline{M}_n$ then, $(x_n)_{n \in \mathbb{N}}$ converges to a solution to Problem* (9)*.*

3.5. Remark

Other convex optimization algorithms can be envisaged to solve Problem (9). On the one hand, it is possible to replace the Dykstra's inner loop by iterations of the Douglas-Rachford algorithm [6]. On the other hand, a minimization strategy relying on the Douglas-Rachford algorithm can be employed where an inner loop consists of iterations of the forward-backward algorithm [12]. In our simulations however, it appeared that the algorithm we propose in this paper is more effective in terms of convergence speed than these two alternatives.

4. SIMULATION RESULTS

The objective of the presented experiment is to recover a microscopy image from a degraded observation z generated according to Model (1). The degradation is obtained by applying a 5×5 positive uniform blur with $\|T\| = 1$. To restore this image, a tight frame version of the dual-tree transform (DTT) proposed in [16] ($\nu = 2$) using Alkin and Caglar 4-band filter bank has been employed. Strictly convex non-differentiable potential functions ϕ_k are chosen, of the form $\omega_k|\,.\,|^{p_k} + \chi_k|\,.\,|$ where $(\omega_k, \chi_k) \in \left]0, +\infty\right[^2$ and $p_k \in \{4/3, 3/2, 2\}$. We have used the algorithm proposed in Section 3.4 with $x_0 = P_C z$ (see (7) for the definition of C) and $\gamma_n = 0.995\ \theta^{-1}$. Visual results are shown in the following figure.



Original Degraded, $\alpha = 0.1$ Restored, $\theta = 0.1$

In Table 1, we provide a quantitative comparison between the regularized Expectation-Maximization algorithm proposed in [7], the Haar-Fisz transform [4], the Anscombe approach in [10] and the proposed method for different values of the quadratic extension parameter θ.

α		0.01	0.05	0.1	1
EM-Reg. [7]		10.7	13.0	14.1	17.8
Haar-Fisz - DTT [4]		8.69	13.1	14.8	18.2
Anscombe - DTT [10]		11.9	15.4	16.4	**19.1**
Quadratic extension DTT	$\theta = 0.001$	**13.6**	**15.6**	16.3	0.00
	$\theta = 0.1$	13.6	15.6	**16.5**	**19.1**

Table 1. SNR for the microscopy image. For quadratic extension method, iteration number $<$500 when $\theta = 0.001$ and $\simeq 1000$ when $\theta = 0.1$.

As mentioned in Section 2.4, a large value of θ allows us to improve the restoration performance (closer approximation g_θ of g), whereas a too small value may yield poorly restored images. However, the step-size γ_n being inversely proportional to θ, a compromise has to be done between the convergence rate of the algorithm and the accuracy to the model. It can be noticed that a significant gain is obtained in comparison with other methods except for the case when $\alpha = 1$ for which the Anscombe approach gives similar results.

5. CONCLUSIONS

A new restoration algorithm has been proposed for wavelet-based restoration in the presence of Poisson noise. This algorithm allows us to minimize a wide class of convex objective functions where the fidelity term is directly related to the Poisson likelihood. A quadratic extension is necessary to ensure the gradient Lipschitz property of the smooth term. Although not thoroughly discussed in this paper, this extension is also beneficial to the improvement of the convergence rate.

6. REFERENCES

[1] C. A. Bouman and K. Sauer, "A unified approach to statistical tomography using coordinate descent optimization," *IEEE Trans. on Image Proc.*, vol. 5, no. 3, pp. 480–492, Mar. 1996.

[2] 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*, vol. 27, no. 7, pp. 943–959, July 2008.

[3] F. J. Anscombe, "The transformation, of Poisson, binomial and negative-binomial data," *Biometrika*, vol. 35, no. 3/4, pp. 246–254, Dec. 1948.

[4] P. Fryzlewicz and G. P. Nason, "A Haar-Fisz algorithm for poisson intensity estimation," *Journal of Computational and Graphical Statistics*, vol. 13, pp. 621–638, 2004.

[5] S. Sardy, A. Antoniadis, and P. Tseng, "Automatic smoothing with wavelets for a wide class of distributions," *Journal on Computational and Graphical Statistics*, vol. 13, pp. 399–421, 2004.

[6] 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.*, vol. 1, no. 4, pp. 564–574, Dec. 2007.

[7] C. L. Byrne, "Iterative image reconstruction algorithms based on cross-entropy minimization," *IEEE Trans. on Image Proc.*, vol. 2, no. 1, pp. 96–103, Jan. 1993.

[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*, vol. 69, pp. 260–266, 2006.

[9] T. D. Capricelli and P. L. Combettes, "A convex programming algorithm for noisy discrete tomoraphy," in *Advances in Discrete Tomography and Its Applications*, pp. 207–226. Boston, MA: Birkhäuser, G. T. Herman and A. Kuba edition, 2007.

[10] 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.*, Las Vegas, Mar. 30-Apr. 4, 2008, pp. 761–764.

[11] N. Pustelnik, C. Chaux, and J.-C. Pesquet, "A constrained forward-backward algorithm for image recovery problems," in *Proc. Eur. Sig. and Image Proc. Conference*, Lausanne, Switzerland, Aug. 25-29, 2008, 5p.

[12] C. Chaux, P. L. Combettes, J.-C. Pesquet, and V. R. Wajs, "A variational formulation for frame-based inverse problems," *Inverse Problems*, vol. 23, pp. 1495–1518, June 2007.

[13] J. A. Fessler, "Hybrid poisson/polynomial objective functions for tomographic image reconstruction from transmission scans," *IEEE Trans. on Image Proc.*, vol. 4, pp. 1439–1450, 1995.

[14] C. Chaux, J.-C. Pesquet, and N. Pustelnik, "Nested iterative algorithms for convex constrained image recovery problems," *http://arxiv.org/abs/0806.3920*, submitted.

[15] H. H. Bauschke and P. L. Combettes, "A Dykstra-like algorithm for two monotone operators," *Pacific Journal of Optimization*, vol. 4, pp. 383–391, Sept. 2008.

[16] C. Chaux, L. Duval, and J.-C. Pesquet, "Image analysis using a dual-tree M-band wavelet transform," *IEEE Trans. on Image Proc.*, vol. 15, no. 8, pp. 2397–2412, Aug. 2006.