

# A Dictionary Learning Approach for Poisson Image Deblurring

Liyan Ma, Lionel Moisan, Jian Yu and Tiejong Zeng\*

**Abstract**—The restoration of images corrupted by blur and Poisson noise is a key issue in medical and biological image processing. While most existing methods are based on variational models, generally derived from a Maximum A Posteriori (MAP) formulation, recently sparse representations of images have shown to be efficient approaches for image recovery. Following this idea, we propose in this paper a model containing three terms: a patch-based sparse representation prior over a learned dictionary, the pixel-based total variation regularization term and a data-fidelity term capturing the statistics of Poisson noise. The resulting optimization problem can be solved by an alternating minimization technique combined with variable splitting. Extensive experimental results suggest that in terms of visual quality, PSNR value and the method noise, the proposed algorithm outperforms state-of-the-art methods.

**Index Terms**—Deblurring, dictionary learning, patch-based approach, Poisson noise, total variation.

## I. INTRODUCTION

Image degradation is unavoidable in real applications, in particular in biomedical imaging, microscopy, astronomical imaging, where low-intensity signals are frequently encountered. Among the most important degradations in these fields are noise, in particular the unavoidable Poisson noise due to the quantum nature of light, and blur, at least caused by the natural extent of the point spread function of the imaging device. Other causes of blur, like defocus or motion blur, are also frequently encountered. This makes image restoration, which aims at recovering a high-quality image from its degraded observation, a very important topic in imaging science in general, and in these fields in particular [1]–[7].

Mathematically, the degraded observed image  $f \in \mathbb{R}^{m \times n}$  can be written as

$$f = \mathcal{P}(Hu), \quad (1)$$

where  $u \in \mathbb{R}^{m \times n}$  denotes the unknown ideal image, the known linear operator  $H : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$  stands for blur,

Manuscript received August 08, 2012; revised February 22, 2013; accepted March 21, 2013. This work was supported in part by RGC of Hong Kong (211710, 211911), the National Natural Science Foundation of China (11271049, 60905028, 61033013, 61105056, 81230086), the National 863 Project (2012AA02A609), the Ph.D. Programs Foundation of Ministry of Education of China (20120009110006), the Fundamental Research Funds for the Central Universities (2011JBM030, 2013JBZ005) and HKBU FRGs.

L. Ma and J. Yu are with School of Computer and Information Technology, Beijing Jiaotong University, Beijing, China. (E-mail: {08112068,jianyuy}@bjtu.edu.cn).

L. Moisan is with Université Paris Descartes, MAP5 (CNRS UMR 8145), Sorbonne Paris Cité, France. (E-mail: lionel.moisan@parisdescartes.fr).

\*T. Zeng is with Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong, China. (E-mail: zeng@hkbu.edu.hk).

Copyright (c) 2010 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org).

and  $\mathcal{P}$  denotes the effect of noise. Gaussian noise, impulse noise, Poisson noise and multiplicative noise are typical instances that have been studied in the literature. When  $H$  is the identity matrix, the problem reduces to image denoising.

As the linear operator  $H$  is usually ill-conditioned, image deblurring is a classical ill-posed inverse problem. Regularity conditions are then required to restrict the solution space [8]–[11]. A widely-used regularization criterion in image processing is total variation (TV) [10], which is known to well preserve edges in images. However, TV regularization is also known to over-smooth textured regions, which may cause the loss of important details. Other widely-used approaches are wavelets and frames [12]–[14], which provide a multiscale and sparse representation of images, hence breaking the ill-posedness of the initial inverse problem. These methods better preserve textures than TV regularization and interesting connections between wavelet-based and TV-based restoration methods were established by Chambolle et al. [15] and later by Steidl et al. [16] and Cai et al. [17]. In 2002, Malgouyres [18] and Candes [19] proposed a hybrid restoration method combining TV and wavelets. Although wavelet-based methods provide good results in many applications, they are based on a fixed dictionary, independent of the image content. As a result, these methods may fail to capture the distinctive characteristics of the processed image, which limits their performances. To overcome this limitation and better take into account the structure of the processed image, patch-based approaches have been proposed and widely adopted in image processing in the past few years.

Many patch-based sparse priors have been studied for image restoration. One approach exploits a heavy-tailed gradient distribution of natural images [20]. Elad and Aharon [21] proposed an effective denoising method (called K-SVD) based on a sparse and redundant representation: their algorithm first learns an optimal over-complete dictionary from the observed noisy image patches, and then recovers each image patch via a linear combination of only a few atoms in the learned dictionary. Another remarkable denoising method relying on a sparse representation is the BM3D algorithm proposed by Dabov et al. [22]. Similar patches are first stacked into 3D arrays, and then jointly denoised using collaborative filtering in the 3D transformed domain. The BM3D method is very efficient, and does not need to find an explicit dictionary. Recently, Zoran and Weiss [23] proposed a simple Gaussian Mixture prior learned from a set of natural images, then used this prior for denoising or deblurring purposes. Although the restoration results obtained by Zoran and Weiss are better than many other sparse prior based methods, the computational cost

of the stage corresponding to the learning of the Gaussian Mixture prior is very high.

Most of the above-mentioned methods are designed to work with Gaussian noise. Unfortunately, Gaussian distributions provide only a limited approximation in most real applications. Applying these methods directly for image deblurring under other kinds of noise is not very efficient in general, because different kinds of noise may affect the image in totally different ways. In particular, Poisson noise, that we consider in the present paper, is signal dependent, as the variance of the noise is, in each pixel, proportional to the intensity of the signal. More precisely, assuming as usual that the values of the observed image  $f$  at the locations  $\{i\}$  are independent, we can write

$$P(f|Hu) = \prod_i \frac{e^{-(Hu)_i} ((Hu)_i)^{f_i}}{f_i!}. \quad (2)$$

Lots of methods have been proposed for Poisson noise removal. Major contributions fall in the following three categories: (a) simply recover the image using a method designed for Gaussian noise removal; (b) transform Poisson noise into near-Gaussian noise by applying an appropriate transform to the noisy image (called a variance-stabilizing transform, VST), then process the transformed image with an algorithm designed for Gaussian noise removal, and finally apply to it the inverse transform in order to get the result [1], [24], [25]; (c) remove Poisson noise directly via a data-fidelity term derived from Poisson noise statistics [14], [26]–[28]. The present contribution belongs to the third category. Let us also mention recent works dealing with mixed Poisson-Gaussian noise removal [1], [24], [29].

Using the Bayesian framework in the case of Poisson noise, Le et al. [27] proposed to minimize

$$\|\nabla u\|_1 + \lambda \langle Hu - f \log Hu, 1 \rangle, \quad (3)$$

where the first term is the classical discrete TV regularization, that is, the  $l_1$  norm of the gradient  $\nabla u = (\partial_x u, \partial_y u)^T$  whose components  $\partial_x u$  and  $\partial_y u$  are finite differences that estimate the partial derivatives of the image. The second term is the data-fidelity term, derived from the log-likelihood of the Poisson distribution. This functional is non-quadratic and causes several difficulties. Le et al. proposed to minimize (3) by using a traditional gradient descent, which is slow. More efficient methods were proposed later: Sawatzky et al. [30] proposed an EM-TV algorithm; Chaux et al. [31] proposed a nested iterative algorithm; Setzer et al. [28] employed the split Bregman technique [32]; Figueiredo et al. [14] used the alternating direction method of multiplier to solve (3) or a frame-based version of (3).

Based on the developments mentioned above and inspired by the K-SVD algorithm for Gaussian noise removal, we here propose a new model, involving a sparse representation over a learned dictionary, to deblur images corrupted by Poisson noise. Although Xiao et al. [33] studied Poisson noise removal via a learned dictionary, they did not consider the blur which, as we saw earlier, cannot be avoided in many real applications. Also, some recent works studied image deblurring under Gaussian noise via a learned dictionary [34]–[36], but as far

as we have known, our work is the first one that deals with Poissonian image deblurring via a learned dictionary.

The paper is organized as follows. In Section II, we briefly review the main principle of K-SVD algorithm. In Section III, we detail the proposed model and propose an algorithm to solve it. In Section IV, some numerical experiments are presented to demonstrate the performance of our model. Other applications, computational time and possible further improvements are then discussed in Section V. Finally, Section VI concludes the paper.

## II. IMAGE DENOISING USING A PATCH-BASED SPARSE REPRESENTATION PRIOR

Choosing a good prior (or equivalently, a good regularization term) is the key of solving most ill-posed inverse problems in Image Processing and Computer Vision. The approach based on a learned dictionary has achieved very promising results in many applications since it was introduced by Olshausen and Field [37] in 1996. In the present paper, we adopt the K-SVD method, proposed by Elad et al. [21], [38], to learn a good sparse representation prior. It is a flexible method, that can be adapted to different tasks, and it has the advantage of providing an explicit dictionary to represent image patches. We now briefly recall the K-SVD method for Gaussian noise removal.

The basic assumption of the K-SVD method is that image patches admit a sparse representation. More precisely, each image patch (of size  $\sqrt{N} \times \sqrt{N}$ , converted to a single column vector  $x \in \mathbb{R}^N$ ) can be compactly represented as a linear combination of a small subset of patches (atoms), taken from a dictionary  $D \in \mathbb{R}^{N \times K}$ . Denoting by

$$\|x\|_p = \left( \sum_{i=1}^m |x_i|^p \right)^{\frac{1}{p}}$$

the classical  $l_p$  norm in Euclidean space for  $p \in [1, \infty)$ , the above assumption can be formally transposed into the minimization problem

$$\hat{\alpha} = \arg \min_{\alpha \in \mathbb{R}^K} \|\alpha\|_0, \quad s.t. \|D\alpha - x\|_2 \leq \varepsilon, \quad (4)$$

where  $\varepsilon$  is zero or a small positive constant, and the “zero-norm”  $\|\alpha\|_0 := \#\{i \mid 1 \leq i \leq K, \alpha_i \neq 0\}$  counts the number of nonzero entries in the vector  $\alpha$ . Under Gaussian noise assumption, the observed noisy image is  $g = u + b$ , where  $u$  is the ideal (non-observed) image to recover and  $b \sim \mathcal{N}(0, \sigma^2 I)$  is the Gaussian noise term. In the following, we shall assume that  $g$  (and  $u$ ) are written as column vectors of  $\mathbb{R}^{m \times n}$  indexed by

$$\mathcal{A} = \{1, 2, \dots, m\} \times \{1, 2, \dots, n\},$$

and the subset of  $\mathcal{A}$  where  $\sqrt{N} \times \sqrt{N}$  patches can be placed will be denoted by

$$\mathcal{A}_s = \{1, 2, \dots, m - \sqrt{N} + 1\} \times \{1, 2, \dots, n - \sqrt{N} + 1\}.$$

Following the sparse representation assumption (4), Elad et al. [21] proposed to achieve Gaussian noise removal with the

variational model

$$\min_{\{\alpha_{ij}\}, u, D} \sum_{(i,j) \in \mathcal{A}_s} \mu_{ij} \|\alpha_{ij}\|_0 + \sum_{(i,j) \in \mathcal{A}_s} \|D\alpha_{ij} - R_{ij}u\|_2^2 + \lambda \|u - g\|_2^2, \quad (5)$$

where the binary matrix  $R_{ij} \in \mathbb{R}^{N \times mn}$  corresponds to the extraction of a  $\sqrt{N} \times \sqrt{N}$  block from the image at location  $(i, j)$ , so that  $R_{ij}u \in \mathbb{R}^N$ . The hidden parameters  $\mu_{ij} > 0$  depend on the optimization procedure. The first two terms in (5) correspond to the above-mentioned sparsity assumption (for some dictionary  $D \in \mathbb{R}^{N \times K}$ ), while the last term controls data fidelity (for the statistics of Gaussian noise, that correspond to a squared  $l_2$  norm), weighted by a positive parameter  $\lambda$ .

The choice of the dictionary  $D$  has a significant impact on the performance of the K-SVD algorithm. Experimental results in [21] show that learning the dictionary from the noise image leads to better performances in most cases. To get the solution of (5), Elad et al. minimized (5) with respect to  $D$ ,  $\alpha_{ij}$  and  $u$  separately:

- Solving for  $D$  given  $\{\alpha_{ij}\}, u$ : Elad et al. update one atom at a time in  $D$  and its corresponding coefficients using SVD decomposition of a residual matrix. Details can be found in [21], [38].
- Solving for  $\{\alpha_{ij}\}$  given  $D, u$  requires to consider

$$\min_{\{\alpha_{ij}\}} \sum_{(i,j) \in \mathcal{A}_s} \mu_{ij} \|\alpha_{ij}\|_0 + \sum_{(i,j) \in \mathcal{A}_s} \|D\alpha_{ij} - R_{ij}u\|_2^2. \quad (6)$$

In general, this optimization problem is difficult to solve. Aharon et al. [38] used the orthogonal matching pursuit (OMP) method [39] to efficiently get a solution.

- Solving for  $u$  given  $D, \{\alpha_{ij}\}$ : The recovered image is a solution of

$$\min_u \sum_{(i,j) \in \mathcal{A}_s} \|D\alpha_{ij} - R_{ij}u\|_2^2 + \lambda \|u - g\|_2^2. \quad (7)$$

This minimization problem has a closed-form solution: taking the derivative of (7) with respect to the vector  $u$ , setting to 0 and solving the resulting equation yields

$$\hat{u} = (\lambda I + \sum_{i,j} R_{ij}^T R_{ij})^{-1} (\lambda g + \sum_{i,j} R_{ij}^T D \alpha_{ij}).$$

The K-SVD algorithm provides very good results for Gaussian noise removal, and it has been later generalized to handle color image denoising, demosaicking, inpainting and image sequence denoising [40], [41]. In the next section, we present a way to adapt this sparse representation prior to the case of Poisson image deblurring.

### III. PROPOSED ALGORITHM

We first describe a new variational model for Poisson image deblurring, then present an algorithm to find the solution by solving the associated minimization problem. (See Table I for a list of the notations used in this paper.)

TABLE I  
NOTATIONS USED THROUGH THE PAPER.

$f$	blurry and noisy image
$g$	noisy image
$u$	the unknown ideal image
$H$	blur kernel
$\mathcal{P}$	the effect of noise
$P(\cdot)$	probability
$\nabla$	the gradient operator
$\lambda$	the balanced parameter for the data-fidelity term
$\sqrt{N} \times \sqrt{N}$	patch size
$\mathcal{A}_s$	set of patches' locations
$D$	dictionary
$\alpha_{ij}$	the coefficients for the patch located at $(i, j)$
$R_{ij}$	the extraction of a patch from the image at location $(i, j)$
$\mu_{ij}$	hidden parameters
$\hat{u}$	the estimation of $u$
$\eta$	the balanced parameter for the TV term
$p, q, w$	the auxiliary variables introduced in the splitting method
$\eta_1, \gamma, \beta$	positive parameters penalize the distance between $u, \nabla, Hu$ and their corresponding auxiliary variables
$\nabla^*$	the complex conjugate transpose of $\nabla$
$N_O$	the iterations of the outer loop
$N_I$	the iterations of the inner loop
$r_\beta, r_\gamma, r_{\eta_1}$	growth rates for parameters $\beta, \gamma, \eta_1$

#### A. Proposed model

Images encountered in real applications are structured data that present lots of repeated patterns, in particular edges, smooth regions, and textures. This is probably why methods incorporating sparse and adaptive patch priors have exhibited very good performances. In the proposed model, we use the sparse representation prior (4) proposed by Elad et al. [21], [38] as a regularization term, and the data-fidelity term of (3) to model blur and Poisson noise. Moreover, to overcome the artifacts sometimes caused by patch-based priors in deblurring tasks (see, e.g., the artifacts encountered in smooth regions in [34], [36] when deblurring under Gaussian noise), we add a TV regularization term. Thus, the discrete variational model we propose for Poisson image deblurring writes

$$\min_{\{\alpha_{ij}\}, u, D} \sum_{(i,j) \in \mathcal{A}_s} \mu_{ij} \|\alpha_{ij}\|_0 + \sum_{(i,j) \in \mathcal{A}_s} \|D\alpha_{ij} - R_{ij}u\|_2^2 + \eta \|\nabla u\|_1 + \lambda \langle Hu - f \log Hu, 1 \rangle, \quad (8)$$

where  $\eta$  and  $\lambda$  are positive constants that balance the different terms. Notations  $\mathcal{A}_s, \mu_{ij}, \alpha_{ij}, D$  and  $R_{ij}$  are those of (5). This model mixes two priors: one is patch-based and it contributes to recover textures, while the other one is pixel-based and helps favoring local smoothness while keeping sharp edges. The last term in (8) is the data-fidelity term, which measures the “distance” between the recovered image and the observed one according to Poisson statistics.

The proposed model (8) raises some important issues: it is non convex (because of the product between the unknowns  $D$  and  $\alpha_{ij}$ ), the TV regularization term is non-differentiable, and the data-fidelity term, because of the log function, is much less easy to handle than, say, a classical squared  $l_2$  norm. Fortunately, these difficulties can be overcome: given  $D$ , the function to minimize with respect to  $\{\alpha_{ij}\}$  and  $u$  is convex, so that there is hope to find a good minimization algorithm by minimizing this convex function in an inner loop, while

optimizing with respect to  $D$  in an outer loop. Moreover, there exists lots of methods to handle non-differentiable functions [32], [42]–[45], so that the TV term is not really a problem. Last, the variable splitting method [32], [42], [45] can be used to overcome the non-quadratic form of the data-fidelity term. Before describing the algorithm that we will use to solve (8), we introduce in the next section the variable splitting method.

### B. Variable Splitting Method

The main idea driving the variable splitting method is that by introducing appropriate auxiliary variables, the minimization of a sum of two functionals involving the same variable can be achieved by means of separate minimizations. To be more precise, let us consider the unconstrained optimization problem

$$\min_x F_1(Ax) + F_2(x), \quad (9)$$

where  $F_1$  and  $F_2$  are two functions and  $A$  is a linear operator. By introducing an auxiliary variable  $d$ , Problem (9) can be rewritten as the constrained problem

$$\min_{d,x} F_1(d) + F_2(x) \quad s.t. \quad Ax = d. \quad (10)$$

We can see that Problem (10) is equivalent to Problem (9) in the feasible set  $\{(d, x) \mid Ax = d\}$ . Now we can relax the constrained problem (10) to an unconstrained one, namely

$$\min_{d,x} F_1(d) + \frac{\beta}{2} \|Ax - d\|_2^2 + F_2(x). \quad (11)$$

The connection between (11) and (10) is simple: when the parameter  $\beta$  is sent to  $+\infty$ , Problem (11) goes back to (10), or, equivalently (9). But in general, solving (11) is much easier than solving (9). Thanks to this variable splitting method, many efficient algorithms have been designed to solve minimization problems involving a non-differentiable TV regularization term [14], [32], [42], [45], [46].

### C. The full algorithm for Poisson image deblurring

Let us now return to the optimization problem (8). Using the variable splitting method presented above, we transform (8) into

$$\begin{aligned} \min_{\{\alpha_{ij}\}, u, D, p, q, w} & \sum_{(i,j) \in \mathcal{A}_s} \mu_{ij} \|\alpha_{ij}\|_0 + \sum_{(i,j) \in \mathcal{A}_s} \|D\alpha_{ij} - R_{ij}p\|_2^2 \\ & + \frac{\beta}{2} \|p - u\|_2^2 + \eta \|q\|_1 + \frac{\eta_1}{2} \|\nabla u - q\|_2^2 \\ & + \lambda \langle w - f \log w, 1 \rangle + \frac{\gamma}{2} \|w - Hu\|_2^2, \end{aligned} \quad (12)$$

where  $p \in \mathbb{R}^{mn}$ ,  $q \in \mathbb{R}^{mn} \times \mathbb{R}^{mn}$ ,  $w \in \mathbb{R}^{mn}$  are three auxiliary variables, and  $\eta_1$ ,  $\gamma$ ,  $\beta$  are three positive numerical parameters, that can be chosen rather large in practice. The new minimization problem (12) is easier to solve than (8). Like many algorithms [43], [45], we can use an alternating minimization algorithm to solve (12). This procedure involves several sub-problems:

1) Solving for  $\{\alpha_{ij}\}$  given  $u, D$ : For  $(i, j) \in \mathcal{A}_s$ , the minimization of (12) with respect to  $\alpha_{ij}$  amounts to

$$\min_{\alpha_{ij}} \mu_{ij} \|\alpha_{ij}\|_0 + \|D\alpha_{ij} - R_{ij}p\|_2^2.$$

This is exactly the sub-optimization problem (6), already present in the K-SVD algorithm. To solve it, we use the OMP method [39] as in the K-SVD algorithm.

2) Solving for  $p$  given  $u, D$ : The minimization of (12) with respect to  $p$  becomes

$$\min_p \sum_{(i,j) \in \mathcal{A}_s} \|D\alpha_{ij} - R_{ij}p\|_2^2 + \frac{\beta}{2} \|p - u\|_2^2. \quad (13)$$

This is a least squares problem, similar to (7). Its closed-form solution is

$$p = (\beta I + 2 \sum_{(i,j) \in \mathcal{A}_s} R_{ij}^T R_{ij})^{-1} (\beta u + 2 \sum_{(i,j) \in \mathcal{A}_s} R_{ij}^T D\alpha_{ij}). \quad (14)$$

3) Solving for  $q$  given  $u, D$ : The minimization of (12) with respect to  $q$  boils down to

$$\min_q \eta \|q\|_1 + \frac{\eta_1}{2} \|\nabla u - q\|_2^2.$$

This is an  $l_1$ -regularized least squares problem. It can be solved explicitly by pointwise soft-thresholding (see also [13], [14], [45]):

$$q = \text{shrink} \left( \nabla u, \frac{\eta}{\eta_1} \right), \quad (15)$$

where the shrinkage operator is defined by

$$\text{shrink}(t, \tau) = \max(\|t\| - \tau, 0) \frac{t}{\|t\|}.$$

4) Solving for  $w$  given  $u, D$ : The minimization of (12) with respect to  $w$  leads to

$$\min_w \lambda \langle w - f \log w, 1 \rangle + \frac{\gamma}{2} \|w - Hu\|_2^2, \quad (16)$$

whose explicit solution is

$$w = \frac{Hu - \frac{\lambda}{\gamma} + \sqrt{\left(Hu - \frac{\lambda}{\gamma}\right)^2 + 4\frac{\lambda f}{\gamma}}}{2}. \quad (17)$$

5) Solving for  $u$  given  $\{\alpha_{ij}\}, p, q, w, D$ : The minimization of (8) with respect to  $u$  is equivalent to

$$\min_u \frac{\beta}{2} \|p - u\|_2^2 + \frac{\eta_1}{2} \|\nabla u - q\|_2^2 + \frac{\gamma}{2} \|w - Hu\|_2^2.$$

This is again a least squares problem, whose closed-form solution is

$$u = (\beta + \eta_1 \nabla^* \nabla + \gamma H^T H)^{-1} (\beta p + \eta \nabla^* q + \gamma H^T w), \quad (18)$$

where the complex conjugate transpose  $\nabla^*$  of  $\nabla$  is defined by

$$\forall q = [q_1, q_2]^T \in U = \mathbb{R}^{mn} \times \mathbb{R}^{mn}, \quad \nabla^* q = \partial_x^* q_1 + \partial_y^* q_2.$$

The products by  $H$  and  $H^T$  in (18) can be efficiently computed by FFT algorithm.

As we saw in Section III-B, Model (12) is equivalent to Model (8) when the parameters  $\eta_1$ ,  $\gamma$ ,  $\beta$  tend to  $+\infty$ . Therefore, we should set these three parameters to large values, so that the solution of (12) remains close to the solution of (8). However, if we use too large values, numerical problems occur in the minimization process (see [47]). This is why, inspired by the implementation of the FTVd method [45], we start with



small values for the three parameters and gradually increase them along the iterations, thereby forcing the convergence to the solution of Model (8).

The dictionary learning procedure has a crucial influence on the quality of the recovered image. Since the degraded image lacks textures and details, we use the estimated image  $u$  to learn the dictionary  $D$ . However, updating  $D$  each time  $u$  changes would be computationally too expensive, so we just update  $D$  in an outer loop. In this paper, we use the method in [21] to learn the dictionary. The details of the proposed algorithm are summarized in Algorithm 1.

---

**Algorithm 1** Proposed algorithm for Poisson image deblurring

---

**Initialization:** Set  $u = f$ ,  $D = DCT$ . Choose  $\eta_1$ ,  $\lambda$ ,  $\gamma$ ,  $\eta$  and  $\beta$ , the number of iterations for the outer ( $N_O$ ) and inner ( $N_I$ ) loops, the growth rates  $r_\beta$ ,  $r_\gamma$ ,  $r_{\eta_1}$

**Output:**  $u$

**for**  $iter_O = 1$  to  $N_O$  **do**

**for**  $iter_I = 1$  to  $N_I$  **do**

    update  $\alpha_{i,j}$  using OMP;

    update  $p$  using (14);

    update  $q$  using (15);

    update  $w$  using (17);

    update  $u$  using (18);

**end for**

  update Dictionary  $D$ ;

$\beta = \beta \cdot r_\beta$ ;

$\gamma = \gamma \cdot r_\gamma$ ;

$\eta_1 = \eta_1 \cdot r_{\eta_1}$ ;

**end for**

---

#### IV. EXPERIMENTAL RESULTS

In this section, we evaluate the performance of the proposed algorithm and compare it to the TV-based algorithm [28] (denoted by PIDSplit+) and to the frame-based algorithm [14] (denoted by PIDAL-FA). We first present experiments on MRI and Fluorescence Microscopy images, then provide experiments on natural images to further study the behavior of these methods.

We consider two different blur kernels: a  $9 \times 9$  Gaussian blur with standard deviation 1 (generated by the Matlab function 'fspecial'), and a  $5 \times 5$  uniform blur. Three peak intensities are considered ( $Peak = 255, 600, 1000$ ), corresponding to different Poisson noise levels. To obtain the degraded image, we first scale the original image by  $Peak/I_{max}$ , where  $I_{max}$  is the maximal value of the original image, and then convolve it with the appropriate blur kernel, before with simulate Poisson noise on the blurred image (Matlab function 'poissrnd').

Periodic boundary conditions are adopted, so that the Discrete Fourier Transform (DFT) can be used to compute the convolution operator. The quality of the restoration results is compared quantitatively by means of the peak signal-to-noise ratio (PSNR), which is defined (as usual) by

$$PSNR = 20 \log_{10} \frac{Peak}{\frac{1}{mn} \|u^* - u\|_2}, \quad (19)$$

where  $u^*$  is the restored image and  $u$  is the true original image.

The stopping criterion we adopt is widely used [14], [45]: we stop when the relative difference between two successive iterates of the restored image fall beyond a fixed threshold  $\varepsilon$ , that is, when

$$\frac{\|u^{k+1} - u^k\|_2}{\|u^{k+1}\|_2} < \varepsilon. \quad (20)$$

We stop the inner loop of Algorithm 1 when  $iter_I = N_I$  or (20) is satisfied with  $\varepsilon = 10^{-5}$ . For PIDSplit+, we use the stopping criterion suggested in [28] and set the error bound to be 0.5. We set  $\varepsilon = 10^{-4}$  for PIDAL-FA. For these settings, all the algorithms reach convergence in most cases.

For PIDSplit+, we set  $\gamma = 10/\lambda$ , and the parameter  $\lambda$  in the regularization term is tuned empirically to perform well.

For PIDAL-FA, we set the parameter  $\mu$  as suggested in [14], and the parameter  $\tau$  in the regularization term is tuned empirically to perform well. As suggested in [14], we use a redundant Haar frame for the Cameraman image and Daubechies-4 for the other images<sup>1</sup>.

For the proposed algorithm, we set the parameters of K-SVD<sup>2</sup> as follows:  $J = 30$  (since this is enough to ensure the convergence), patches of size  $4 \times 4$ , and a dictionary of size  $16 \times 256$ . We also set  $\eta = 0.1$ ,  $N_O = 12$ ,  $N_I = 60$ , and the initialization for  $\eta_1$ ,  $\beta$  and  $\gamma$  are 0.01, 10 and 50 respectively (excepted for Barbara image in Section IV-B, for which we take  $\gamma = 20$ ). Growth rates are  $r_\beta = r_\gamma = 2$  and  $r_{\eta_1} = 1.5$ , and the remaining parameter  $\lambda$  in the data-fidelity term is tuned empirically to achieve good performances.

#### A. Experiments on Medical Images

To compare the effects of PIDSplit+ and PIDAL-FA algorithms to ours, we realized synthetic experiments (blur + noise, then deblurring) on the following gray-level images coming from the biomedical field<sup>3</sup>: *Ankle* ( $512 \times 512$ ), *Brain* ( $256 \times 256$ ), *Mouseintestine4* ( $256 \times 256$ ) and *Neck* ( $256 \times 256$ ). These images are presented in Fig.1.

In Fig. 2 is reported a detail of the results obtained by processing a corrupted version of *Ankle* image (Gaussian blur and Poisson noise with peak intensity 600). We can observe that the TV-based PIDSplit+ method tends to oversmooth the image and exaggerate the width of fine structures. It also presents a blocky aspect along some edges. The frame-based PIDAL-FA method preserves more details, but introduces some artifacts. Comparatively, our method reaches a higher PSNR value and achieves better visual quality. The same conclusions hold for *Brain* image corrupted with a uniform blur and Poisson noise (see Fig.3). For the latter experiment, we show in Fig.4 the final dictionary learned at the last iteration of the proposed method.

<sup>1</sup>The frame is computed via the Rice Wavelet toolbox: <http://dsp.rice.edu/software/rice-wavelet-toolbox>

<sup>2</sup>When implementing the K-SVD denoising algorithm, we use the code provided by Rubinstein on his website: <http://www.cs.technion.ac.il/~ron-rubin/software.html>

<sup>3</sup>We downloaded these images from the following websites: <http://sehati.org/index/patientresources/diagnosticprocedures/mri.html>, <http://www.cedars-sinai.edu/Patients/Programs-and-Services/Imaging-Center/For-Patients/Exams-by-Procedure/MRI/Lower-Extremities-Leg.aspx>, <http://www.microscopyu.com/articles/livecellimaging/index.html>, <http://osc-ortho.com/blog/tag/open-mri/>

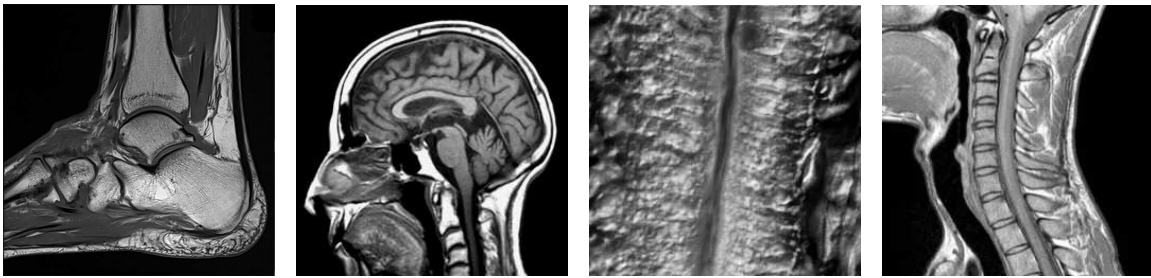


Fig. 1. Original images. From left to right: *Ankle, Brain, Mouseintestine4, Neck.*

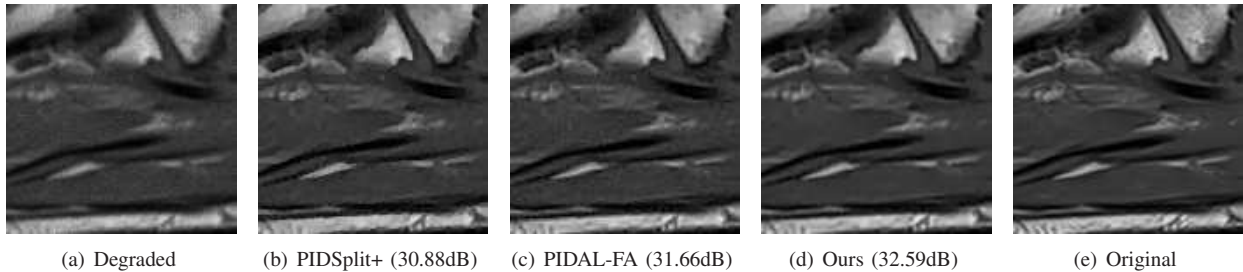


Fig. 2. Results of different methods on *Ankle* image (detail) corrupted by a Gaussian blur ( $\sigma = 1$ ) and Poisson noise with peak intensity 600.

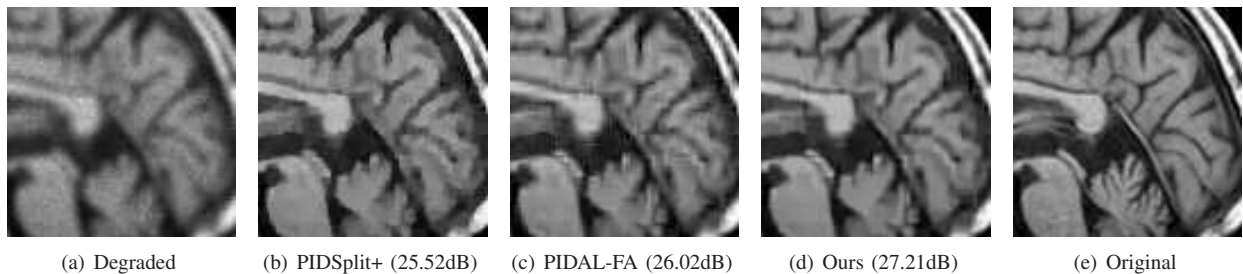


Fig. 3. Results of different methods on *Brain* image (detail) corrupted by a  $5 \times 5$  uniform blur and Poisson noise with peak intensity 600.

Note that our dictionary method assumes high repetition of features in the image. This is not always true in medical images as minor irregularities may sometimes be of high importance, as pointed out by [48]. In order to handle this problem, we have reduced the dictionary patch size from  $16 \times 16$  (in the original K-SVD paper [21]) to  $4 \times 4$ . It is then interesting to see that the dictionary shown in Fig.4 contains some typical details of the original *Brain* image, thus showing the ability of the algorithm to recover details from the degraded image. The spotted atoms there should be useful to represent those minor irregularities in medical images (see spots in Fig. 11).

More systematic results are presented in Table II.

Besides PSNR evaluations, we also provide results of the *Method Noise* proposed in [49] on the *Brain* image, adaptive to Poisson noise. Denote  $u$  as the underlying image (possibly with some noise), and  $D_h u$  the image restoration result with parameter(s)  $h$ . Inspired by [25], [48], [49], the *Method Noise* (or the stabilized noise residual) is then defined as

$$\mathcal{A}(u) - \mathcal{A}((D_h u)), \quad (21)$$

where  $\mathcal{A}$  is the identity operator for Gaussian noise or the Anscombe transform (see [25]) for Poisson noise for variance stabilization. The importance of *Method Noise* was discussed

in [48]. Evidently, few textures/structure should be present in the difference in (21). Since we are mainly dealing with Poisson noise, which is signal dependent, we apply the Anscombe transform on both  $u$  and  $D_h u$ , leading to an approximate Gaussian noise. The results of *Method Noise* are reported in Fig. 5 where all the *Method Noise* images are drawn with the same grayscale range. Note that for fair comparisons, we tune the parameters in different methods until the MSE between  $u$  and  $D_h u$  reaches a predefined value. Clearly, our method produces the best result with the least image structure information removed.

The parameters settings corresponding to the different peak intensities ( $Peak = 1000, 600, 255$ ) are, respectively: for the PIDSplit+ method,  $\lambda = 0.01, 0.012, 0.02$  (Gaussian blur) and  $\lambda = 0.005, 0.007, 0.015$  (uniform blur); for the PIDAL-FA method,  $\tau = 0.01, 0.015, 0.025$  (Gaussian blur) and  $\tau = 0.008, 0.01, 0.02$  (uniform blur); for the method we propose, we took  $\lambda = 5000, 4000, 3000$  (Gaussian blur) and  $\lambda = 8000, 7000, 5000$  (uniform blur).

## B. Experiments on Natural Images

To further compare the performance of the three methods, we also ran experiments on classical gray-level images: *Bar-*

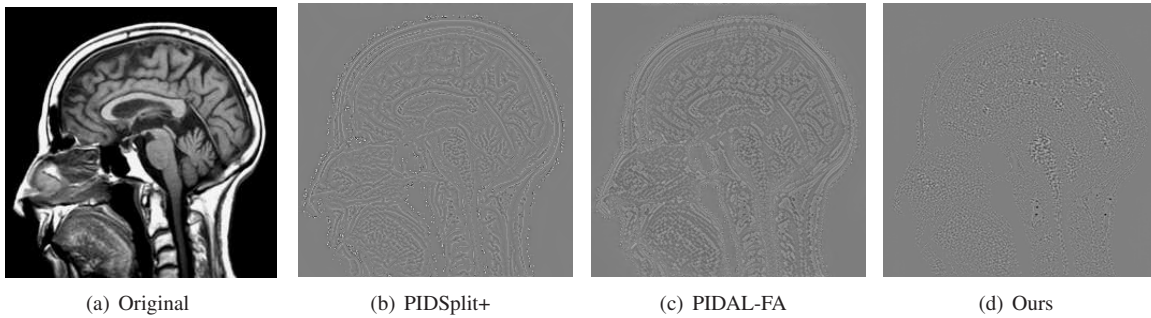


Fig. 5. Method Noise of different methods on the *Brain* image.

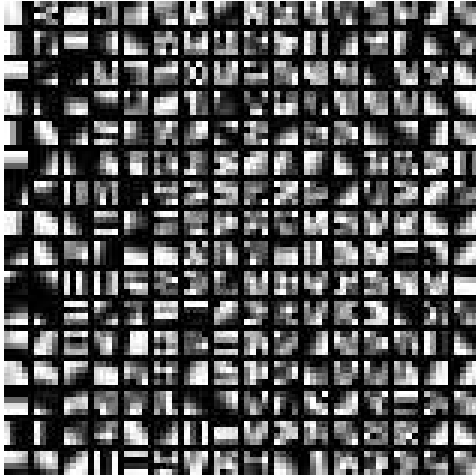


Fig. 4. Dictionary learned by our method for the experiment of Fig.3.

TABLE II  
PSNR(DB) VALUES FOR PIDSPLIT+ [28], PIDAL-FA [14] AND OUR ALGORITHM FOR THE TEST IMAGES OF FIG.1 CORRUPTED BY A  $9 \times 9$  GAUSSIAN BLUR WITH STANDARD DEVIATION 1 OR BY A  $5 \times 5$  UNIFORM BLUR, AND POISSON NOISE WITH VARIOUS PEAK INTENSITIES ( $Peak = 1000, 600, 255$ ).

Images/Peak	Gaussian blur			Uniform blur		
	[28]	[14]	Ours	[28]	[14]	Ours
Ankle/1000	31.54	32.45	<b>33.56</b>	28.74	29.16	<b>29.77</b>
Ankle/600	30.88	31.66	<b>32.59</b>	28.15	28.47	<b>29.22</b>
Ankle/255	29.63	30.13	<b>30.84</b>	27.40	27.73	<b>28.08</b>
Brain/1000	29.22	29.93	<b>31.66</b>	26.22	26.68	<b>27.41</b>
Brain/600	28.42	29.01	<b>30.79</b>	25.52	26.02	<b>27.21</b>
Brain/255	27.08	27.50	<b>29.38</b>	24.40	24.85	<b>26.14</b>
Mou./1000	29.78	30.48	<b>31.91</b>	27.01	27.59	<b>28.26</b>
Mou./600	29.16	29.77	<b>31.04</b>	26.37	26.90	<b>27.50</b>
Mou./255	27.80	28.26	<b>29.22</b>	25.43	26.05	<b>26.07</b>
Neck/1000	29.36	29.91	<b>30.91</b>	27.01	27.52	<b>27.71</b>
Neck/600	28.78	29.23	<b>30.23</b>	26.33	26.88	<b>27.28</b>
Neck/255	27.64	27.98	<b>28.85</b>	25.39	25.87	<b>26.24</b>
Aver./1000	29.98	30.69	<b>32.01</b>	27.25	27.74	<b>28.29</b>
Aver./600	29.06	29.92	<b>31.16</b>	26.59	27.07	<b>27.80</b>
Aver./255	28.04	28.47	<b>29.57</b>	25.66	26.13	<b>26.63</b>

*bara* ( $512 \times 512$ ), *Cameraman* ( $256 \times 256$ ), *Lena* ( $512 \times 512$ ) and *Man* ( $512 \times 512$ ). The original images are shown in Fig.6.

In Fig. 7 is reported a detail of the restored images obtained from *Barbara* image (degraded with a Gaussian blur and

Poisson noise with peak intensity 600). The PIDSplit+ method recovers less details than PIDAL-FA, which is itself a little less contrasted and detailed than our result, as confirmed by the PSNR values. By employing the multiscale structure information, the PIDAL-FA method can better recover textures made of repeated patterns than the PIDSplit+ method, but for the *Man* image, which contains many different irregular textures, the superiority of PIDAL-FA over PIDSplit+ vanishes. For *Cameraman* image (Fig.9), the level of recovered details and the PSNR are in accordance: our method outperforms PIDAL-FA, which is itself a little above PIDSplit+. Another observation we can make is that like all TV-based methods, PIDSplit+ produces images that suffer from the staircase effect (piecewise constant regions sometimes delimited by artificial boundaries, instead of smoothly-varying gray values). This artifact does not appear with our method, that permits smooth transitions as each patch is represented by a linear combination of atoms taken from the learned dictionary.

More systematic results are presented in Table III. The parameters settings corresponding to the different peak intensities ( $Peak = 1000, 600, 255$ ) are, respectively: for the PIDSplit+ method,  $\lambda = 0.008, 0.015, 0.025$  (Gaussian blur) and  $\lambda = 0.006, 0.01, 0.02$  (uniform blur); for the PIDAL-FA method on *Barbara* image  $\tau = 0.008, 0.015, 0.025$  (Gaussian blur) and  $\tau = 0.008, 0.01, 0.02$  (uniform blur), while for other images  $\tau = 0.015, 0.02, 0.03$  (Gaussian blur) and  $\tau = 0.01, 0.015, 0.025$  (uniform blur); for the proposed method, we took  $\lambda = 7000, 6000, 4000$  (Gaussian blur) and  $\lambda = 8000, 7000, 5000$  (uniform blur).

## V. DISCUSSIONS

In this section, we discuss several topics associated to the present work: blind deblurring, a possible adaptation to image deblurring under multiplicative noise, the effect of the parameters and the convergence of the algorithm, computation time aspects, and suggestions for further improvements.

### A. Blind deblurring

Until now, we assume that the blur kernel is known. However, this is not possible in real applications. Estimating kernels, which is a very difficult problem, is still an undergoing topic. Here, we just do some attempts to handle blind deblurring under Poisson noise. Since almost all the existing methods designed for estimating kernels are under





Fig. 6. Original images. From left to right: *Barbara*, *Cameraman*, *Lena*, *Man*.



Fig. 7. Results of different methods on *Barbara* image corrupted by the Gaussian blur and Poisson noise with peak intensity 600.

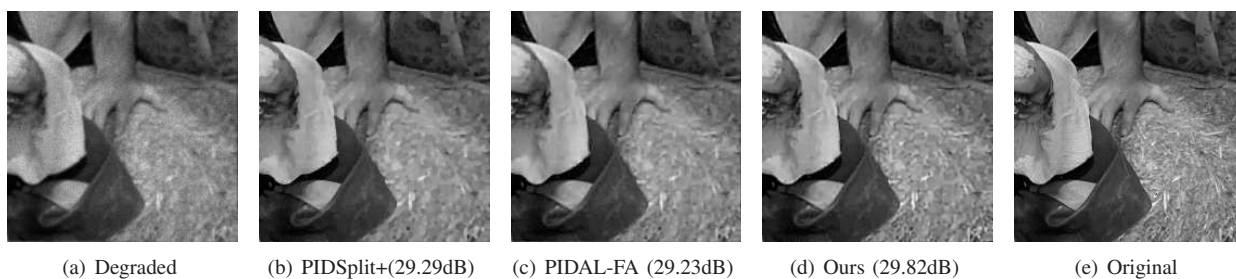


Fig. 8. Results of different methods on *Man* image corrupted by the Gaussian blur and Poisson noise with peak intensity 600.

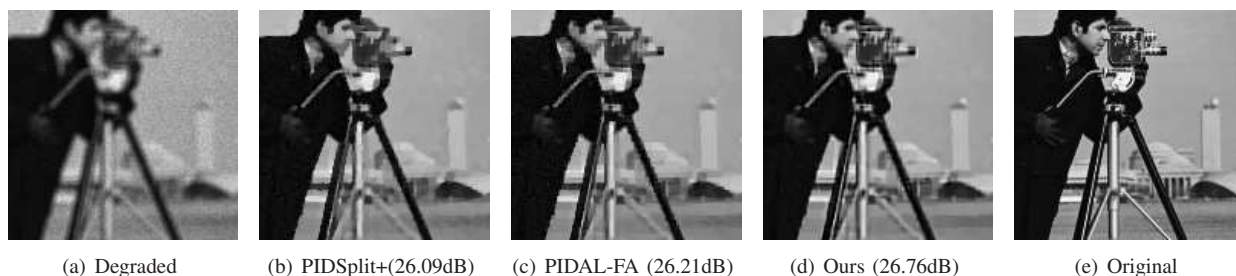


Fig. 9. Results of different methods on *Cameraman* image corrupted by the Uniform blur and Poisson noise with peak intensity 600.

the Gaussian assumption of noise, we handle the Poisson noise via a variance stabilizing transformation (VST)(see [25] for example) leading to an approximate normal distribution. Then we use the algorithm proposed by Krishnan et al. [50] to estimate the kernel. Having the kernels, we can use the proposed method to recover the image.

First, we present an experiment on the synthetic image: degraded *Brain* image corrupted by a Gaussian blur and Poisson noise with peak intensity 600. Fig. 10 presents the results. The size of the original kernel is 9. The size of the estimated kernel is set to be 15. We can observe that the proposed method achieves higher restored quality than the compared algorithms. However, the resulting PSNR value

is much lower than that obtained by deblurring with known kernel.

Second, we present experiments on real images<sup>4</sup> with Poisson noise. The sizes of the estimated kernels are set to 15. We tune parameters empirically to obtain the best visual results for all the methods. In Fig. 11, we test our method on a CT image. We can observe that the proposed method achieves higher restored quality, with noise suppressed and blur reduced. In Fig. 12, we provide the *Method Noise* on the CT image to test the effects of different methods. The *Method Noise*

<sup>4</sup>We downloaded these two images from the DICOM Sample Image Sets: <http://www.osirix-viewer.com/datasets/>



TABLE III

PSNR(DB) VALUES FOR PIDSPLIT+ [28], PIDAL-FA [14] AND OUR ALGORITHM FOR THE TEST IMAGES OF FIG.6 CORRUPTED BY A  $9 \times 9$  GAUSSIAN BLUR WITH STANDARD DEVIATION 1 OR BY A  $5 \times 5$  UNIFORM BLUR, AND POISSON NOISE WITH VARIOUS PEAK INTENSITIES ( $Peak = 1000, 600, 255$ ).

Images/Peak	Gaussian blur			Uniform blur		
	[28]	[14]	Ours	[28]	[14]	Ours
Bar./1000	24.98	25.46	<b>25.67</b>	24.24	24.52	<b>24.65</b>
Bar./600	24.51	24.93	<b>25.29</b>	23.95	24.23	<b>24.40</b>
Bar./255	24.07	24.34	<b>24.60</b>	23.57	23.79	<b>23.84</b>
Cam./1000	28.33	28.38	<b>28.77</b>	26.72	26.86	<b>26.97</b>
Cam./600	27.54	27.84	<b>28.40</b>	26.09	26.21	<b>26.76</b>
Cam./255	26.63	26.79	<b>27.52</b>	25.23	25.31	<b>26.06</b>
Lena/1000	32.70	32.98	<b>33.70</b>	30.92	31.25	<b>31.75</b>
Lena/600	32.27	32.38	<b>33.06</b>	30.44	30.70	<b>31.19</b>
Lena/255	31.10	31.15	<b>31.68</b>	29.65	29.75	<b>29.94</b>
Man/1000	29.92	29.78	<b>30.35</b>	28.12	28.17	<b>28.46</b>
Man/600	29.29	29.23	<b>29.82</b>	27.66	27.65	<b>27.99</b>
Man/255	28.34	28.25	<b>28.73</b>	26.90	26.90	<b>27.12</b>
Aver./1000	28.98	29.15	<b>29.62</b>	27.50	27.70	<b>27.96</b>
Aver./600	28.40	28.60	<b>29.14</b>	27.04	27.20	<b>27.59</b>
Aver./255	27.54	27.63	<b>28.13</b>	26.34	26.44	<b>26.74</b>

images are still drawn with the same grayscale range. It is clear that our method removes the least details. Notice the ‘‘Y’’ shaped bright area in the original image. Since this part has a constant grayscale value, it shall not be altered by the deblurring algorithms. This is clearly the case in our method. In Fig. 13, we test our method on a cardiac CT image. We can observe that the restored image by the proposed method is much sharper than the original image. This indicates that the motion blur in the original image is successfully reduced. Meanwhile, we can also observe that the proposed method preserves more details than the other deblurring methods.

The results obtained by blind deblurring under Poisson noise are not satisfactory. There are seldom works handling this topic. Jiang et al. [51] proposed a blind deblurring algorithm to improve spiral CT image resolution in the case of Gaussian blurring. In Computer Vision, many algorithms have been proposed to estimate the blur kernel from images acquired under Gaussian noise (see [52]–[54] and references therein). Although we try to use those methods to estimate the blur kernel by means of a variance stabilizing transformation (VST). However, we think generalizing these works to estimate a blur kernel under Poisson noise would lead to much better recovered results. Furthermore, the blur kernel may be spatial-varying in real applications, this makes the problem more harder to handle.

### B. Deblurring under Multiplicative Noise

Multiplicative noise appears in several applications, such as laser imaging, ultrasound imaging, and synthetic aperture radar (SAR) [55]–[59]. In these cases, the image formation process (including blur) can be written

$$f = (Hu)M_n,$$

where  $M_n$  represents a random variable (multiplicative noise) with mean one. As shown in [57], Model (3) is also appropriate

for multiplicative noise removal when the multiplicative noise  $M_n$  follows the Gamma distribution with probability density function

$$P(x; \theta, L) = \frac{1}{\theta^L \Gamma(L)} x^{L-1} e^{-\frac{x}{\theta}} \quad \text{for } x \geq 0,$$

where  $\Gamma$  is the Euler Gamma function and  $\theta$  and  $L$  denote the scale and shape parameters in the Gamma distribution, respectively. Models of multiplicative noise usually assume that the means equals one, so that  $L\theta = 1$ .

Here we just give a simple example to show that our model, though designed for Poisson noise, can exhibit good performances for image deblurring under multiplicative Gamma noise. We compare five methods: the three methods considered earlier for Poisson image deblurring (including ours), plus two models explicitly designed for image deblurring under multiplicative noise: the AA model [55] and the RLO model [56]. As we can observe in Fig.14, the result obtained with our method leads to a better image quality and a higher PSNR than the other methods considered.

Note that in the case of multiplicative noise removal, bias correction is important [60], [61]. In our experiments, applying a bias correction improves the PSNR by 0.1 to 0.2 dB. This is why, in order to provide a fair comparison, we applied the same projection step proposed in [60] to the results of all algorithms. We refer to [60], [61] for more information on bias correction.

### C. The effect of the parameters and the convergence of the algorithm

Here we study the effect of the parameters in the energy function(12). Actually, most of them can be fixed in all the experiments. The parameter  $\eta$  controls the effect of the TV regularization. When there are some artifacts in the recovered image, one may take larger  $\eta$ . In fact, the influence of the TV regularization also depends on the ratio of the value of  $\eta$  and  $\lambda$ . Thus, we can fix other parameters, and study the sensitivity of the recovered results to the choice of the parameters  $\gamma$  and  $\lambda$ . Fig. 15 gives the PSNR values of the resulting images under certain choice of the two parameters. As expected, we found that the choice of  $\lambda$  is dependent on the noise level, and the choice of  $\gamma$  is dependent on the choice of  $\lambda$ . However, the sensitivity of the results to the choice of  $\gamma$  is smaller.

Fig. 16 presents the objective function values of the proposed method on *Brain* image corrupted by the Gaussian blur and Poisson noise with peak intensity 600. We can observe that, as iterations progress, the objective function value of the proposed method converges

### D. About the Computation Time

Although providing better recovered images, the proposed method is computationally heavier: with our implementation, processing a  $512 \times 512$  image takes around ten minutes, while PIDSplit+ and PIDAL-FA need less than one minute (all experiments realized using Matlab 7.12 (R2011a) on a PC equipped with a 2.70GHz CPU).

The computation time cost of our approach could be reduced in at least three ways: (1) using a better computer or even

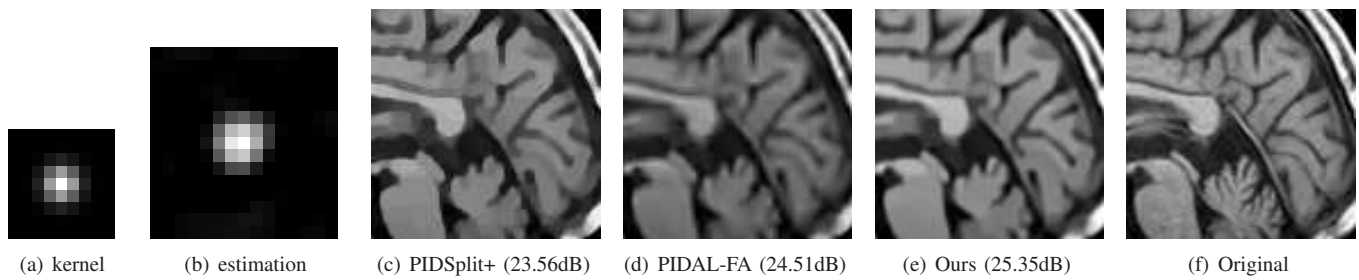


Fig. 10. Results of different methods on *Brain* image corrupted by a Gaussian blur and Poisson noise with peak intensity 600.

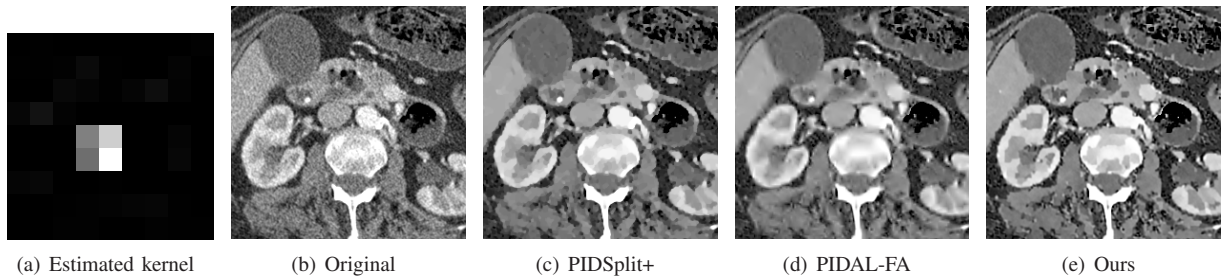


Fig. 11. Results of different methods on *CT* image.

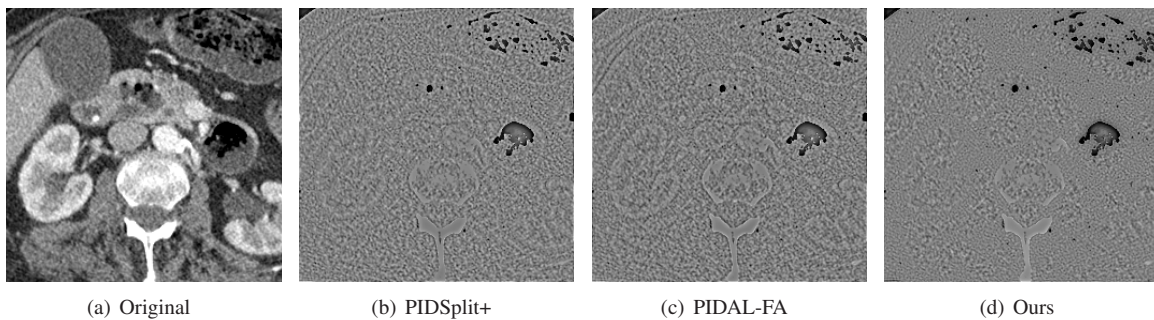


Fig. 12. *Method Noise of different methods on the CT image.*

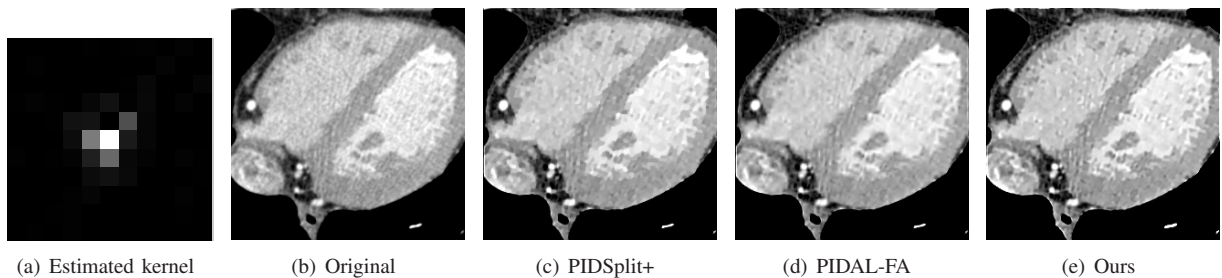


Fig. 13. Results of different methods on a cardiac *CT* image.

clusters since ours for experiments is not so good; (2) applying some paralleling techniques to compute the most time costing step (6); (3) stopping the outside/inside iterations timely once the satisfactory resulting image is obtained. Currently we set  $N_I = 60$  and  $N_O = 20$  which are largely unnecessary in many cases.

#### E. How to further improve the proposed method

The proposed method could be improved in different ways. A more powerful dictionary learning method could be used, incorporating for example the “non-local” information as in

[62], thus taking advantage of the self-similarity of images [63], [64] to constrain similar patches to share similar or identical dictionary elements in their sparse representation. Another possibility would be to incorporate the multiscale structure of images as in [65], since the size of texture patterns varies among different images, or even inside a single image (actually, for some test images considered in Section IV, we got better results when using different patch sizes, but we chose to use the same patch size for simplification). Last, other patch-based sparse priors could also improve the results, in particular the Gaussian mixture prior used in [23] or the image content based heavy-tailed gradient distribution of [20].

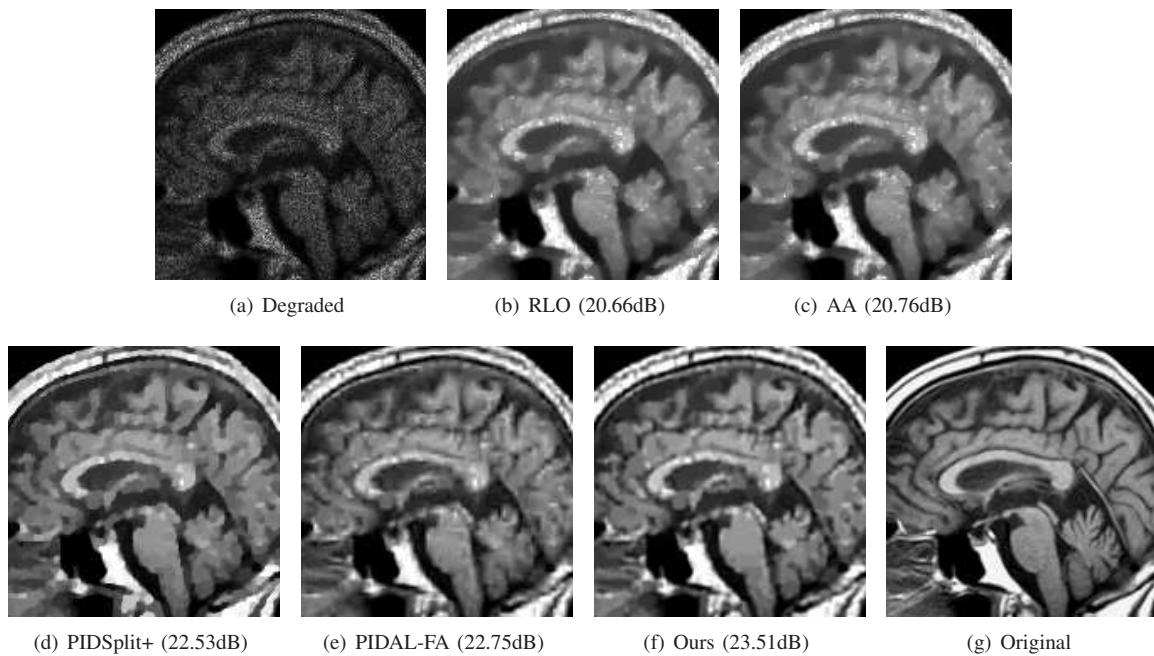


Fig. 14. Results obtained with different methods on *Brain* image corrupted by Gaussian blur ( $\sigma = 1$ ) and multiplicative noise ( $L = 10$ ).

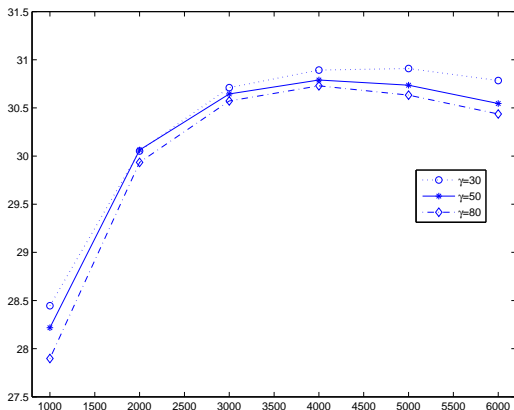


Fig. 15. Resulting PSNR values of our method on *Brain* image corrupted by the Gaussian blur and Poisson noise with peak intensity 600 when increasing the values of  $\lambda$  for different  $\gamma$ .

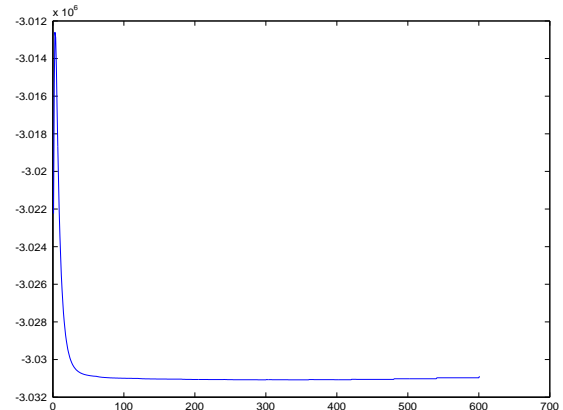


Fig. 16. Resulting objective function values of our method on *Brain* image corrupted by the Gaussian blur and Poisson noise with peak intensity 600 as iterations progress.

## VI. CONCLUSION

In this paper, we proposed a new model to recover images suffering from blur and Poisson noise. It is based on the mixture of two priors: a patch-based sparse representation prior over a learned dictionary, and the pixel-based TV regularization. Compared to TV-based and frame-based methods, the proposed algorithm leads to higher PSNR values and improves the quality of restored images, though requiring a higher computational cost.

Among future research directions, we think it could be worth exploring strategies to reduce the computation time. We also plan to work on possible extensions of the present method to different frameworks such as color images, video sequences, or other non-Gaussian noise models (like Rician noise, that appears in MRI [48], [66]). Designing a blind deblurring algorithm by embedding a blur kernel estimation

procedure in the present approach would also be an interesting development and open new possibilities of applications.

## VII. ACKNOWLEDGEMENTS

The authors also thank the anonymous reviewers for their extremely useful suggestions for improving the quality of the paper.

## REFERENCES

- [1] J. Boulanger, C. Kervrann, P. Bouthemy, P. Elbau, J. Sibarita, and J. Salamero, "Patch-based nonlocal functional for denoising fluorescence microscopy image sequences," *IEEE Trans. Medical Imaging*, vol. 29, no. 2, 2010.
- [2] R. Chan and K. Chen, "Multilevel algorithms for a poisson noise removal model with total variation regularization," *Int. J. Comput. Math.*, vol. 84, pp. 1183–1198, 2007.
- [3] G. Hagberg, G. Zito, F. Patria, and J. Sanes, "Improved detection of event-related functional mri signals using probability functions," *NeuroImage*, vol. 2, pp. 1193–1205, 2001.



- [4] J. Ollinger and J. Fessler, "Positron emission tomography," *IEEE Signal Process. Mag.*, vol. 14, no. 1, pp. 43–55, 1997.
- [5] J. Miao and D. Wilson, "Selective evaluation of noise, blur, and aliasing artifacts in fast mri reconstructions using a weighted perceptual difference model: Case-pdm," in *Proc. SPIE 7263, 72631N*, 2009.
- [6] E. Rollano-Hijarrubia, R. Manniesing, and W. Niessen, "Selective deblurring for improved calcification visualization and quantification in carotid ct angiography: Validation using micro-ct," *IEEE Trans. Medical Imaging*, vol. 28, no. 3, 2009.
- [7] G. Wang, M. Vannier, M. Skinner, M. Cavalcanti, and G. Harding, "Spiral ct image deblurring for cochlear implantation," *IEEE Trans. Medical Imaging*, vol. 17, no. 2, 1998.
- [8] N. Hurley and S. Rickard, "Comparing measures of sparsity," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4723–4741, 2009.
- [9] G. Peyre, S. Bougleux, and L. Cohen, "Non-local regularization of inverse problems," *Inverse Probl. Imag.*, vol. 5, pp. 511–530, 2011.
- [10] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.
- [11] A. Tikhonov and V. Arsenin, *Solution of Ill-Posed Problems*. PWinston, Washington, DC, 1977.
- [12] D. Donoho and I. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.
- [13] I. Daubechies, M. Defriese, and C. Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, pp. 1413–1457, 2004.
- [14] M. Figueiredo and J. Bioucas-Dias, "restoration of poisson images using alternating direction optimization," *IEEE Trans. Image Process.*, vol. 19, pp. 3133–3145, 2010.
- [15] A. Chambolle, R. DeVore, N.-Y. Lee, and B. Lucier, "Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage," *IEEE Trans. Image Process.*, vol. 7, pp. 319–335, 1998.
- [16] G. Steidl, J. Weickert, T. Brox, P. Mrzek, and M. Welk, "On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and sides," *SIAM J. Numer. Anal.*, vol. 42, pp. 686–713, 2004.
- [17] J. Cai, B. Dong, S. Osher, and Z. Shen, "Image restoration: Total variation, wavelet frames, and beyond," *J. Amer. Math. Soc.*, vol. 25, pp. 1033–1089, 2012.
- [18] F. Malgouyres, "Mathematical analysis of a model which combines total variation and wavelets for image restoration," *Journal of information processes*, vol. 2, pp. 1–10, 2002.
- [19] E. Candes and F. Guo, "A new multiscale transforms, minimum total variation synthesis: application to edge-preserving image reconstruction," *Signal Processing*, vol. 82, pp. 1519–1543, 2002.
- [20] T. Cho, N. Joshi, C. Zitnick, S. Kang, R. Szeliski, and W. Freeman, "A content-aware image prior," in *CVPR*, (San Francisco, USA), pp. 169–176, IEEE, 2010.
- [21] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, pp. 3736–3745, 2006.
- [22] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3d transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, pp. 2080–2095, 2007.
- [23] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in *ICCV*, (Barcelona, Spain), pp. 479–486, Nov. 2011.
- [24] M. Mäitalo and A. Foi, "Optimal inversion of the generalized anscombe transformation for poisson-gaussian noise," *preprint*, 2012.
- [25] B. Zhang, M. Fadili, and J. Stack, "Wavelets, ridgelets, and curvelets for poisson noise removal," *IEEE Trans. Image Process.*, vol. 17, no. 7, pp. 1093–1108, 2008.
- [26] M. Carlván and L. Blanc-Feraud, "Sparse poisson noisy image deblurring," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1834–1846, 2012.
- [27] T. Le, R. Chartran, and T. Asaki, "A variational approach to reconstructing images corrupted by poisson noise," *J. Math. Imaging Vis.*, vol. 27, pp. 257–263, 2007.
- [28] S. Setzer, G. Steidl, and T. Teuber, "Deblurring poissonian images by split bregman techniques," *J. Vis. Commun. Image R.*, vol. 21, pp. 193–199, 2010.
- [29] F. Luisier, T. Blu, and M. Unser, "Image denoising in mixed poisson-gaussian noise," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 696–708, 2011.
- [30] A. Sawatzky, C. Brune, F. Wubbeling, T. Kusters, K. Schafers, and M. Burger, "Accurate em-tv algorithm in pet with low snr," in *Proc. IEEE Nuclear Science Symposium Conference Record*, pp. 5133 – 5137, 2008.
- [31] C. Chau, J. Pesquet, and N. Pustelnik, "Nested iterative algorithms for convex constrained image recovery problems," *SIAM J. Imag. Sci.*, vol. 2, pp. 730–762, 2009.
- [32] T. Goldstein and S. Osher, "The split bregman method for l1-regularized problems," *SIAM J. Imag. Sci.*, vol. 2, pp. 323–343, 2009.
- [33] Y. Xiao and T. Zeng, "Poisson noise removal via learned dictionary," in *ICIP*, pp. 1177–1180, IEEE, 2010.
- [34] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1838–1857, 2011.
- [35] Y. Lou, A. Bertozzi, and S. Soatto, "Direct sparse deblurring," *J. Math. Imaging Vis.*, vol. 39, no. 1, pp. 1–12, 2011.
- [36] H. Zhang, J. Yang, Y. Zhang, and T. Huang, "Sparse representation based blind image deblurring," in *ICME*, (Barcelona, Spain), pp. 1–6, IEEE, July 2011.
- [37] B. Olshausen and D. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, pp. 607–609, 1996.
- [38] M. Aharon, M. Elad, and A. Bruckstein, "K-svd: An algorithm for designing of overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, pp. 4311–4322, 2006.
- [39] J. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2231–2242, 2004.
- [40] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. Image Process.*, vol. 17, pp. 53–69, 2008.
- [41] M. Protter and M. Elad, "Image sequence denoising via sparse and redundant representations," *IEEE Trans. Image Process.*, vol. 18, pp. 27–35, 2009.
- [42] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.*, vol. 19, pp. 2345–2356, 2010.
- [43] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, pp. 681–695, 2011.
- [44] A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math. Imag. Vis.*, vol. 20, pp. 89–97, 2004.
- [45] J. Yang, Y. Zhang, and W. Yin, "An efficient tvl1 algorithm for deblurring multichannel images corrupted by impulsive noise," *SIAM J. Sci. Comput.*, vol. 31, pp. 2842–2865, 2009.
- [46] W. Zuo and Z. Lin, "A generalized accelerated proximal gradient approach for total variation-based image restoration," *IEEE Trans. Image Process.*, vol. 20, no. 10, pp. 2748–2759, 2011.
- [47] J. Nocedal and S. Wright, *Numerical Optimization, 2nd Edition*. Springer, 2006.
- [48] Y. Gal, A. Mehnert, A. Bradley, K. McMahan, D. Kennedy, and S. Crozier, "Denoising of dynamic contrast-enhanced mr images using dynamic nonlocal means," *IEEE Trans. Medical Imaging*, vol. 29, no. 2, 2010.
- [49] A. Buades, B. Coll, and J. Morel, "A non-local algorithm for image denoising," *IEEE Computer Society Conference on.*, vol. 2, pp. 60–65, 2005.
- [50] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in *CVPR*, pp. 233–240, 2011.
- [51] M. Jiang, G. Wang, M. Skinner, J. Rubinstein, and M. Vannier, "Blind deblurring of spiral ct images," *IEEE Trans. Medical Imaging*, vol. 22, no. 7, 2003.
- [52] J. Cai, H. Ji, C. Liu, and Z. Shen, "Blind motion deblurring from a single image using sparse approximation," in *CVPR*, pp. 104–111, 2009.
- [53] T. Cho, S. Paris, B. Horn, and W. Freeman, "Blur kernel estimation using the radon transform," in *CVPR*, pp. 241–248, 2011.
- [54] R. Fergus, B. Singh, A. Hertzmann, S. Roweis, and W. Freeman, "Removing camera shake from a single photograph," in *SIGGRAPH*, pp. 783–794, 2006.
- [55] G. Aubert and J. Aujol, "A variational approach to remove multiplicative noise," *SIAM J. Appl. Math.*, vol. 68, pp. 925–946, 2008.
- [56] L. Rudin, P. Lions, and S. Osher, *Geometric Level Set Methods in Imaging, Vision, and Graphics*, ch. Multiplicative denoising and deblurring: theory and algorithms, pp. 103–119. Springer, 2003.
- [57] G. Steidl and T. Teuber, "Removing multiplicative noise by douglas-rachford splitting methods," *J. Math. Imaging Vis.*, vol. 36, pp. 168–184, 2010.
- [58] Y. Huang, L. Moisan, M. Ng, and T. Zeng, "Multiplicative noise removal via a learned dictionary," *IEEE Trans. Image Process.*, vol. 21, no. 11, pp. 4534–4543, 2012.

- [59] Y. Huang, M. Ng, and T. Zeng, "The convex relaxation method on deconvolution model with multiplicative noise," *Commun. Comput. Phys.*, vol. 13, no. 4, pp. 1066–1092, 2013.
- [60] Y. Dong and T. Zeng, "A convex variational model for restoring blurred images with multiplicative noise," *UCLA Cam-report 12-23*, March 2012.
- [61] S. Durand, J. Fadili, and M. Nikolova, "Multiplicative noise removal using l1 fidelity on frame coefficients," *J. Math. Imaging Vis.*, vol. 36, pp. 201–226, 2010.
- [62] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *ICCV*, pp. 2272–2279, 2009.
- [63] A. Buades, B. Coll, and J. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, pp. 490–530, 2005.
- [64] A. Effros and T. Leung, "Texture synthesis by non-parametric sampling," in *7th IEEE Int. Conf. Comput. Vis.*, vol. 2, pp. 1033–1038, 1999.
- [65] J. Mairal, G. Sapiro, and M. Elad, "Learning multiscale sparse representations for image and video restoration," *Multiscale Model. Simul.*, vol. 7, pp. 214–241, 2008.
- [66] S. Awate and R. Whitaker, "Feature-preserving mri denoising: A non-parametric empirical bayes approach," *IEEE Trans. Medical Imaging*, vol. 26, no. 9, 2007.