Complex wavelet regularization for 3D confocal microscopy deconvolution

Mikael Carlavan, Laure Blanc-Féraud

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Complex wavelet regularization for 3D confocal microscopy deconvolution

Mikael Carlavan — Laure Blanc-Féraud

N° 7366 — version 2

initial version August 2010 — revised version October 2010

Vision, Perception and Multimedia Understanding

ISSN 0249-6399 ISRN INRIA/RR--7366--FR+ENG

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
SOPHIA ANTIPOLIS - MÉDITERRANÉE

Complex wavelet regularization for 3D confocal microscopy deconvolution

Mikael Carlavan, Laure Blanc-Féraud

Theme : Vision, Perception and Multimedia Understanding
Perception, Cognition, Interaction
Équipe-Projet Ariana

Rapport de recherche n° 7366 — version 2 — initial version August 2010 —
revised version October 2010 — 39 pages

Abstract: Confocal microscopy is an increasingly popular technique for 3D imaging of biological specimens which gives images with a very good resolution (several tenths of micrometers), even though degraded by both blur and Poisson noise. Deconvolution methods have been proposed to reduce these degradations, some of them being regularized on a Total Variation prior, which gives good results in image restoration but does not allow to retrieve the thin details (including the textures) of the specimens. We first propose here to use instead a wavelet prior based on the Dual-Tree Complex Wavelet transform to retrieve the thin details of the object. As the regularizing prior efficiency also depends on the choice of its regularizing parameter, we secondly propose a method to select the regularizing parameter following a discrepancy principle for Poisson noise. Finally, in order to implement the proposed deconvolution method, we introduce an algorithm based on the Alternating Direction technique which allows to avoid inherent stability problems of the Richardson-Lucy multiplicative algorithm which is widely used in 3D image restoration. We show some results on real and synthetic data, and compare these results to the ones obtained with the Total Variation and the Curvelets priors. We also give preliminary results on a modification of the wavelet transform allowing to deal with the anisotropic sampling of 3D confocal images.

Key-words: 3D confocal microscopy, deconvolution, complex wavelet regularization, discrepancy principle, Alternating Direction technique, Richardson-Lucy algorithm.

Centre de recherche INRIA Sophia Antipolis – Méditerranée
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Régularisation par ondelettes complexes pour la déconvolution en microscopie confocale 3D

Résumé : La microscopie confocale est une méthode de plus en plus utilisée pour l'imagerie 3D de spécimens biologiques offrant des images à une très bonne résolution (quelques dixièmes de micromètres), même si elles sont dégradées à la fois par du flou et par un bruit de Poisson. Des méthodes de déconvolution ont été proposées pour réduire ces dégradations, certaines étant régularisées sur un *a priori* de Variation Totale, qui, même s'il donne de bons résultats en restauration d'images, ne permet pas de récupérer les détails fins (dont les textures) des spécimens. Nous proposons ici d'utiliser un *a priori* de transformée en ondelettes basée sur la transformée en ondelettes "dual-tree" complexes afin de retrouver les détails fins de l'objet. L'efficacité d'un *a priori* de régularisation dépendant également du choix de son paramètre de régularisation, nous proposons dans un second temps une méthode permettant de choisir ce paramètre de régularisation d'après un principe de similitude dans le cas de bruit de Poisson. Enfin, nous proposons un algorithme de restauration basé sur la technique de minimisation par directions alternées qui permet d'éviter les problèmes de stabilité inhérents à l'algorithme multiplicatif de Richardson-Lucy qui est largement utilisé en restauration d'images 3D. Nous montrons des résultats sur des images synthétiques et réelles, et comparons avec les résultats obtenus avec *a priori* de Variation Totale et de transformée en curvelets. Nous donnons également des résultats préliminaires sur une modification de la transformée en ondelettes permettant de prendre en compte l'échantillonnage anisotrope des images 3D confocales.

Mots-clés : Microscopie confocale 3D, déconvolution, régularisation par ondelettes complexes, principe de discrépance, technique des directions alternées, algorithme de Richardson-Lucy.

Acknowledgments

This research work was funded by the ANR DetectFine. The authors gratefully acknowledge Praveen Pankajakshan, Daniele Graziani for several interesting discussions and Pierre Weiss for suggesting us the Alternating Direction Method. We are also grateful to Fanny Pignolé (UMR 6543 CNRS/J.-A. Dieudonné), Luis Almeida (Laboratoire J.-A. Dieudonné) and Stéphane Noselli (UMR 6543 CNRS) for providing the data.

Contents

1 Introduction

This section introduces the confocal microscopy principle, and presents the organisation of this report.

1.1 Confocal microscopy

Confocal microscopy imaging has been introduced by M. Minksy in 1955 [26]. This technique is based on the principle of fluorescence: we "color" some parts (the ones we want to observe) of a specimen with specific dyes (excited by specific wavelength). When we light the specimen with this specific wavelength, the molecules of dyes are excited to a high electronic level. Quickly these molecules will loose this energy and then emit photons (fluorescence phenomena), which pass through the objective, the pinhole, and are then collected by the photomultiplier detector of the microscope (see figure 1 for an illustration of a confocal microscope). The pinhole is a small diaphragm which lets light emitted from the focal point pass through but rejects light from out-of-focus regions of the specimen.

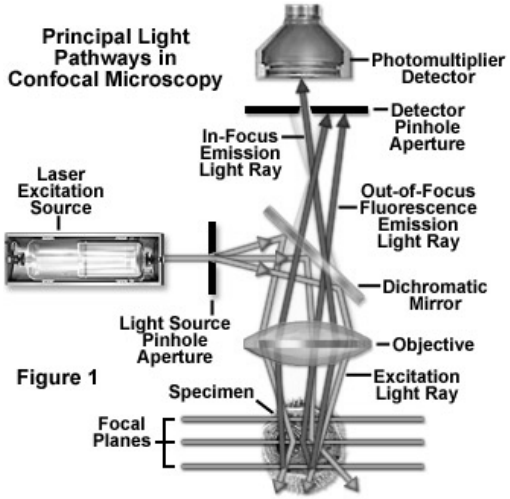


Figure 1: Illustration of a confocal microscope. The laser is the excitation light source. After being reflected by the dichromatic mirror, the light arrives at the objective lens (bottom lens), which focuses it onto a specimen point. If it is fluorescent, it reemits some light. A part of this emitted light is then detected by the photomultiplier (top), located just after the pinhole. From `http://www.microscopyu.com`.

Confocal microscopy imaging offers several advantages over optical imaging such as a small depth-of-field, a reduction of out-of-focus blur and the ability of scanning 3D images. These advantages explain its quick grow in popularity during the last years. However, it suffers from two basic degradations: remaining blur from diffraction and from the out-of-focus light despite the pinhole and

Poisson noise which is due to the weak quantity of light detected.

If we consider a discrete version of a specimen $\mathbf{x} \in \mathbb{R}^n$ (n being the number of voxels of the image) observed as an image $\mathbf{y} \in \mathbb{R}^n$ through an optical system with a Point Spread Function (PSF) h and corrupted by a Poisson noise process $\mathcal{P}$, then the image formation model can be written as [39]:

$$\mathbf{y} = \mathcal{P}(\mathbf{H}\mathbf{x}) \tag{1}$$

where $\mathbf{H} : \mathbb{R}^n \to \mathbb{R}^n$ stands for the matrix notation of the convolution of the PSF h. A good estimation of the PSF h is very important for any non-blind deconvolution algorithm. In this paper, we will use the model presented in [16, 38].

Using a bayesian approach, we want to retrieve the image which maximizes the likelihood probability of (1). This probability can be expressed as:

$$p(\mathbf{y}|\mathbf{x}) = \prod_{\mathbf{i}\in\mathbf{\Omega}} \left(\frac{\left[(\mathbf{H}\mathbf{x})^{\mathbf{y}}\right]_{\mathbf{i}} . \exp\left[-\mathbf{H}\mathbf{x}\right]_{\mathbf{i}}}{\mathbf{y}_{\mathbf{i}}!} \right) \tag{2}$$

Maximizing (2) with respect to $\mathbf{x}$ is equivalent to minimize $-\log p(\mathbf{y}|\mathbf{x})$ i.e. to minimize:

$$J_L(\mathbf{x}, \mathbf{y}) = \mathbf{1}^T (\mathbf{H}\mathbf{x}) - \mathbf{y}^T \log[\mathbf{H}\mathbf{x}] \tag{3}$$

The most used deconvolution algorithm to solve (3) in confocal microscopy may be the Richardson-Lucy (RL) algorithm [24, 33]. This algorithm takes into account Poisson statistics of the photon couting noise, and implicitly imposes positivity constraint on the solution. This is, however, not sufficient to prevent noise amplification during the deconvolution process and this algorithm is usually stopped after an arbitrary number of iterations.
Many authors favour instead the introduction of an explicit prior on the solution to regularize the ill-posed inverse problem and thus minimize a penalized likelihood as in [16, 17, 30]. The authors of [16] have introduced a Total Variation (TV) regularization which smooth homogeneous areas while preserving edges. However small details and textures are not preserved. In [17, 30], a tight-frame synthesis l^1-norm term is used to promote sparsity of the solution in the frame basis. In this work, we study a frame analysis prior in order to regularize the inverse problem, minimizing the l^1-norm of wavelet frame coefficients.

1.2 Contributions of the paper

Our contributions are follows. We introduce a new regularization model for biological images defined by the l^1-norm of a complex wavelet transform in order to model small details and textures. Secondly, we propose a method to select the regularizing parameter which is based on a discrepancy principle for Poisson noise. Finally, we propose to perform the minimization using an algorithm based on the Alternating Direction Method (ADM) instead of the multiplicative Richardson-Lucy algorithm in order to avoid, among other difficulties, stability issue of this algorithm. We give results using TV and analysis frame regularization on biological images.

2 State of art

In this section, we present a state of art of deconvolution methods in confocal microscopy. This section is splitted in three parts. The first one is focused on the priors which have been proposed in confocal microscopy and the second one presents regularizing parameter selection methods for Poisson noise. The last part is devoted to minimization algorithms.

2.1 Priors for confocal miscroscopy

As discussed previously, many works promote the introduction of explicit priors on the solution to regularize the ill-posed inverse problem. Maximizing the a posteriori probability $p(\mathbf{x}|\mathbf{y}) = p(\mathbf{y}|\mathbf{x})\frac{p(\mathbf{x})}{p(\mathbf{y})}$, where $p(\mathbf{x})$ is the prior model on the object given by $p(\mathbf{x}) = \alpha \exp[-\tau J_R(\mathbf{x})]$ (α is a normalization constant and J_R is the regularizing term), is equivalent to:

$$\begin{array}{rl} \arg\min & J(\mathbf{x},\mathbf{y}) := J_L(\mathbf{x},\mathbf{y}) + \tau J_R(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in \mathbb{R}^n \end{array} \tag{4}$$

τ being the regularizing parameter. To the best of our knowledge, the first regularizing term proposed in confocal microscopy was the Tikhonov-Miller regularization given by $J_R(\mathbf{x}) = \|\nabla \mathbf{x}\|_2^2$ [40]. This regularization is efficient to remove noise but its main drawback is that it smooths the edges. To avoid this side effect, the authors of [16] proposed to use instead the l^1-norm of the gradient leading to a well-known regularization in 2D image processing called the Total Variation (TV) [34]. The TV removes the noise while saving the discontinuities but smooths the details of the textures and the corners.
Wavelets priors have been successfully used in 2D image processing to retrieve thin elements including textures ([25] and references therein). But it is only recently that these priors have been introduce in confocal microscopy [10]. These priors assume (and it is actually verified) that images have a compact representation (sparsity) in some wavelet basis. This, in the end, gives a good ability to remove the noise from the image. This sparsity can be forced when a l^1-norm term is used. For example, [10] used a wavelet Haar transform as a prior and [17] proposed to use a decomposition on a dictionnary composed of an undecimated wavelet and a curvelet transform. However, [17] only consider 2D images. On 3D images, using an undecimated wavelet transform is critical as it gives an image which is $7L$ (L being the number of decomposition levels) times higher and is, thus, problematic in term of memory cost. Finally, let us note that [30] showed that these wavelet priors give better results when combined to the Total Variation.

2.2 Regularization parameter selection methods

In most of the deconvolution methods proposed in the literature, the regularization parameter has to be chosen such that it gives the best qualitative results. However, the interpretation of an image may be difficult in biology for example, specially when we use priors which introduce artifacts (such as wavelets priors). To overcome this problem, [17] proposed to use the generalized cross validation (GCV) [22] on the Anscombe transform of the Poisson noise to estimate

the regularization parameter. The authors of [32] extended the recent work of [31] to an unbiaised estimator of the Mean Squared Error (MSE), leading to a pixel-dependent estimator. Thanks to the linearity of the Tikhonov-regularized algorithm, this estimator can be easily implemented. Recently [2] presented three different methods to select the regularization parameter. We recall the one based on a discrepancy principle as the proposed method also follows this idea of discrepancy.

The idea of [2] is to use a Taylor approximation to obtain a quadratic approximation of the term $J_L(\mathbf{x})$ in (3) and to use a discrepancy principle on the approximation. First they showed that if we consider J_L as a function of $\mathbf{x}$ and $\mathbf{y}$, then the Taylor approximation around the exact objects $\bar{\mathbf{x}}$, $\bar{\mathbf{y}} = \mathbf{H}\bar{\mathbf{x}}$ writes:

$$J_L(\mathbf{x}, \mathbf{y}) \simeq J_L(\bar{\mathbf{x}}, \bar{\mathbf{y}}) + J_L^{wls}(\mathbf{x}, \mathbf{y}) \tag{5}$$

with:

$$J_L^{wls}(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \| \left(\mathbf{H}\mathbf{x} - \mathbf{y}\right) / \sqrt{\mathbf{y}} \|_2^2 \tag{6}$$

Using the expected value function E (following the distribution law of $\mathbf{y}$), we can write that around $\bar{\mathbf{x}}$:

$$J_L(\bar{\mathbf{x}}, \mathbf{y}) \simeq J_L(\bar{\mathbf{x}}, \bar{\mathbf{y}}) + E(J_L^{wls}(\bar{\mathbf{x}}, \mathbf{y})) \tag{7}$$

Denoting $\mathbf{x}_\tau$ an estimate of the solution of (4) and using (5), it is reasonable to write:

$$J_L(\mathbf{x}_\tau, \mathbf{y}) \simeq J_L(\bar{\mathbf{x}}, \bar{\mathbf{y}}) + J_L^{wls}(\mathbf{x}_\tau, \mathbf{y}) \tag{8}$$

By combining (7) and (8) we can say that a good value of τ is the one which verifies:

$$J_L^{wls}(\mathbf{x}_\tau, \mathbf{y}) = E(J_L^{wls}(\bar{\mathbf{x}}, \mathbf{y})) \tag{9}$$

[2] showed that $E(J_L^{wls}(\bar{\mathbf{x}}, \mathbf{y}))$ can be well estimated using a common approximation [23]:

$$\mathbf{y} = \mathbf{H}\bar{\mathbf{x}} + \mathbf{e} \tag{10}$$

where $\mathbf{e}$ is a Gaussian random variable with mean 0 and multidimensional variance $\mathbf{y}$. If we set:

$$\mathbf{r}(\mathbf{x}) = \left(\mathbf{H}\mathbf{x} - \mathbf{y}\right) / \sqrt{\mathbf{y}} \tag{11}$$

Then $\mathbf{r}(\bar{\mathbf{x}})$ is a Gaussian random variable with mean 0 and variance $\mathbf{I}$. In this case, a standard result gives:

$$\|\mathbf{r}(\bar{\mathbf{x}})\|_2^2 \sim \chi^2(n) \tag{12}$$

where $\chi^2(n)$ is the chi-square distribution with n degree of freedom which has a mean equal to n. Using this result and (9) we get that a good value of τ verifies:

$$2J_L^{wls}(\mathbf{x}_\tau, \mathbf{y}) \simeq n \tag{13}$$

Finally, the discrepancy principle proposed by [2] to find τ is the following:

$$\tau_{opt} = \begin{array}{cc} \arg\min & \left(2J_L^{wls}(\mathbf{x}_\tau, \mathbf{y}) - n\right)^2 \\ \text{subject to} & \tau \in \mathbb{R}^+ \end{array} \tag{14}$$

2.3 Algorithms for confocal microscopy

The most widely used algorithm in confocal microscopy may be the Richardson-Lucy (RL) algorithm [24, 33]. Using the fact that the PSF is normalized, minimizing (3) leads to the RL algorithm (multiplicative form):

$$\mathbf{x}_{k+1} = \mathbf{x}_k . \left\{ \mathbf{H}^* \left[\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}} \right] \right\} \tag{15}$$

where $\mathbf{H}^*$ denotes the adjoint operator of $\mathbf{H}$. Here, multiplication and division must be understood as point to point operations. This algorithm has two interesting properties. It preserves the number of counts of the original object and has also the property the non-negativity: if the first estimate is positive, then the further estimates stay positive. This algorithm improves the quality of images, however it amplifies the noise after several iterations [41].
Adding the Tikhonov-Miller regularization in the model leads to the following multiplicative algorithm:

$$\mathbf{x}_{k+1} = \frac{\mathbf{x}_k}{1 - 2\tau \mathrm{div}\left(\nabla \mathbf{x}_k\right)} . \left\{ \mathbf{H}^* \left[\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}_k} \right] \right\} \tag{16}$$

where ∇ is the gradient operator and div is the divergence operator (we will use the discretization proposed in [9] for the implementation of these operators). As discussed previously, using the l^2-norm smoothes the edges. The authors of [16] used instead a l^1-norm and obtained the following algorithm:

$$\mathbf{x}_{k+1} = \frac{\mathbf{x}_k}{1 - \tau \mathrm{div}\left(\frac{\nabla \mathbf{x}_k}{|\nabla \mathbf{x}_k|} \right)} . \left\{ \mathbf{H}^* \left[\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}_k} \right] \right\} \tag{17}$$

This algorithm, and more generally algorithms built under the multiplicative form of the Richardson-Lucy algorithm, may suffer from unstability. Several authors proposed to extend "well-known" 2D deconvolution algorithms to 3D confocal microscopy. For example, the authors of [10] proposed to use the "forward-backward" algorithm [11]. However, this algorithm can not be directly used here as the Poisson model leads to solve a problem which does not belong to the class of problems of this algorithm (the obtained criterion to minimize is convex but does not have the Lipschitz gradient property required by [11]). In consequence, the authors of [10] proposed to use a variance stabilizing transform (VST) on the data, the Anscombes [1] transform in that case, such that the Poisson noise is approximated as a Gaussian noise. [17] also proposed to use the Anscombes transform but refined the model such that the VST is taken into account and then solve the problem using an extension of the algorithm [10]. These VST may however not be efficient for images with a weak intensity as in confocal microscopy so [13] proposed a quadratic extension of the Poisson criterion such that "forward-backward" algorithm can be directly used on a Poisson model.

Non-iterative methods can also be used as in [32], where the algorithm used is the Tikhonov-regularized algorithm which leads to an explicit solution. However, the latter consider a model contening both Poisson and Gaussian noise. Thus this model may not be very efficient when considering only Poisson noise

with weak intensity.

Recently, [14] introduced an algorithm which is able to minimize the sum of an arbitrary number of convex functions. [30] showed that this algorithm can be used on the Poisson deconvolution problem. [37] proposed also an efficient algorithm based on split Bregman techniques which really takes into account the Poisson noise statistics. These techniques consist in augmenting the size of the problem by adding several variables and then to solve the problem following each variable. This is actually closely related to the proposed algorithm in this paper, which is based on the Alternating Direction Method (ADM). The ADM has already been proposed for deblurring Poissonian images in [20]. We present here an alternative technique to split the problem which avoids the inner loop present in [20] (in case of Total Variation regularization). A personal communication with one author of [20] revealed however that this inner loop actually accelerates the whole algorithm. It seems that the more we split the problem using the ADM method, the slower becomes the algorithm. This will be subject to further work.

3 Proposed restoration method

In this section we propose to use a complex wavelet transform as a regularizing operator and present a method to select the regularizing parameter for Poisson noise. We finally extend two algorithms to the proposed prior, the RL multiplicative algorthim and the algorithm based on the ADM method.

3.1 Proposed prior

As shown for example in [17], an undecimated wavelet transform and more generally the decomposition on dictionnary using several wavelet transforms clearly improves the quality of the restored image. However this regularization technique for 3D images is really critical as it needs a huge amount of memory. For this reason, we propose here to use the Dual-Tree Complex Wavelet Transform [35] which is a en efficient wavelet transform with a reduced redundancy (8 in 3D). Unlike [17] who use a prior expressed in the transform domain (synthesis prior), we will express our prior in the image domain (analysis prior), as this last one seems to give better results than a synthesis prior (at least for Gaussian noise [7, 18, 36]).

Since several years, wavelet transforms have been used in image processing to restore thin details of complex scenes. Among other shortcomings, usual wavelet transforms suffer from shift variance and from a lack of directional selectivity. To overcome these problems, several authors [12, 35] propose to build a wavelet transform using two trees, the whole transform forming an Hilbert pair (one tree is 90° out of phase of the other). These dual-tree wavelet transforms improve the directionality of real non-redundant wavelet transform with still a low redundancy. We propose to use a recent extension of these dual-tree to the complex case [35]. This Dual-Tree Complex Wavelet (DTCW) transform is *quasi*-invariant by translation and rotation (this property is needed to use a transform as a regularizing operator) and uses two real trees combined to give complex coefficients (illustrated on the figure 2) with a redundancy of 8 in 3D.

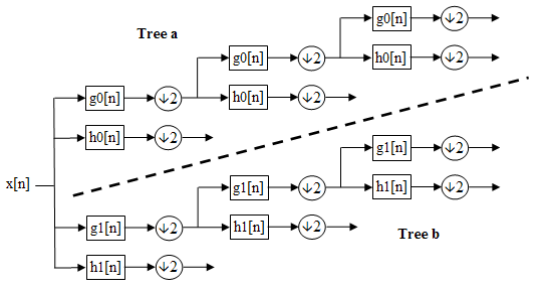


Figure 2: Decomposition scheme of the DTCW transform for an 1D signal. At each level, the details coefficients of each tree give the real and the imaginary part of the complex coefficients.

It has be shown in [35] that thresholding the absolute value of this complex coefficients gives less artifacts that real non-redundant wavelet transforms and thus leads to better restoration results. Moreover, the combination of the two trees gives oriented wavelets which are thus able to analyse oriented objects in image like edges or surfaces. The reason for this behaviour is that the whole wavelet function (denoted by $\psi(u) = \psi_a(u) + j\psi_b(u)$, ψ_a and ψ_b being the wavelets of the trees a and b) is a complex function. Hence, its Fourrier transform is supported on only one side of the frequency plane (unlike real functions which are supported on the two sides of the frequency plane). Then, with a 2D wavelet $\psi_1(u,v) = \psi(u)\psi(v)$, the Fourier transform of $\psi_1(u,v)$ is represented on only one corner of the frequency plane. If the take the real part of this wavelet, we get a symmetric Fourier transform supported on two opposite corners of the frequency plane, giving the -45°oriented wavelet (second wavelet from the left of the first row of the figure 3). The +45°oriented wavelet is obtained by taking $\psi_2(u,v) = \psi(x)\overline{\psi(v)}$, giving a Fourier transform supported on the two other corners (representing by the fifth wavelet of the first row of the same figure). The others wavelets are obtained by considering $\phi(u)\psi(v)$, $\psi(u)\phi(v)$, $\phi(u)\overline{\psi(v)}$ and $\psi(u)\overline{\phi(v)}$.

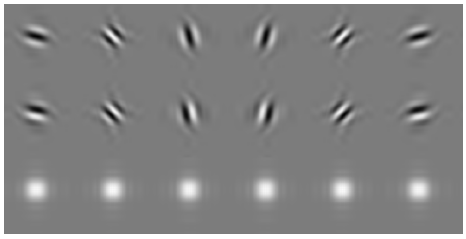


Figure 3: Wavelets of the 2D DTCW transform. The first two rows represent the real and the imaginary part of the wavelet. The third row represents the magnitude of the wavelet.

This wavelet transform has already been sucessfully used to improve the results of the restoration of oriented details, compared to the ones obtained with a Total Variation regularization, in the case of Gaussian noise [8] on 2D images. This wavelet transform has also been proposed in the domain of confocal microscopy in [29] but only for denoising. We will show here that we can include the deconvolution process in the algorithms. To improve the results, we also propose to use a subband-dependent regularization parameter. The proposed regularization writes:

$$J_R(\mathbf{x}) = \sum_{j=1}^{L} \alpha_j \|\mathbf{W}_j \mathbf{x}\|_1 \tag{18}$$

where L is the number of decomposition levels, $\mathbf{W}_j$ is the DTCW decomposition on the level j and α_j is the subband-dependent scale parameter. As we deal with 3D data, the scale of 3D coefficients is decreasing at each scale by a factor $2\sqrt{2}$, thus we will take:

$$\alpha_j = \left(2\sqrt{2}\right)^{-j} \tag{19}$$

Low-pass coefficients are not included in the formulation (18). As there is no reason for these coefficients to be sparse, we set $\alpha_0 = 0$. By integrating the scaling dependance in the transform $\mathbf{W}_j$, (18) can be written as :

$$J_R(\mathbf{x}) = \|\mathbf{W}\mathbf{x}\|_1 \tag{20}$$

$\mathbf{W}$ standing for the whole normalized transform. We will present two algorithms extended to this regularization. We would like also to mention that, usually, images in confocal microscopy do not have the same resolution in the three directions. Some work has been done in this report trying to take into account this difference of resolution, however this has not been included in the whole method as this does not give competitive results (see appendix B).

3.2 Regularizing parameter selection proposed method

The proposed method to select the regularizing parameter is based on the recent work of [42] in which is introduced a discrepancy principle for Poisson noise but only considering denoising. We show in the next lines that this could be extended to the deconvolution. First, let us consider the following function:

$$f(\mathbf{y}_\lambda) = 2\left(\mathbf{y}_\lambda \log\left(\frac{\mathbf{y}_\lambda}{\lambda}\right) + \lambda - \mathbf{y}_\lambda\right) \tag{21}$$

where $\mathbf{y}_\lambda$ is a Poisson random variable with mean λ. Then, [42] showed that:

$$E(f(\mathbf{y}_\lambda)) = 1 + O\left(\frac{1}{\lambda}\right) \tag{22}$$

The considered model in [42] is $\mathbf{y} = \mathcal{P}(\mathbf{x})$. We easily see that this model fits to the function (21) with $\lambda = \mathbf{x}$. So, from this function, [42] defined:

$$J_L^{ber}(\mathbf{x}_\tau, \mathbf{y}) = \frac{2}{n}\left(\mathbf{y}^T \log\left(\frac{\mathbf{y}}{\mathbf{x}_\tau}\right) + \mathbf{1}^T\mathbf{x}_\tau - \mathbf{1}^T\mathbf{y}\right) \tag{23}$$

If one compute this function at the true object $\mathbf{x}$, one get that $J_L^{ber}(\mathbf{x}, \mathbf{y}) \simeq 1$. Thus, [42] proposed to select τ as the one which verifies:

$$\tau_{opt} = \begin{array}{cc} \arg\min & \left(J_L^{ber}(\mathbf{x}_\tau, \mathbf{y}) - 1\right)^2 \\ \text{subject to} & \tau \in \mathbb{R}^+ \end{array} \tag{24}$$

In the case of deconvolution, the statement (22) still holds as we have $\mathbf{y} = \mathcal{P}(\mathbf{H}\mathbf{x})$ and thus $\mathbf{y}$ is a Poisson random variable with mean $\mathbf{H}\mathbf{x}$. In this case $J_L^{ber}(\mathbf{x}_\tau, \mathbf{y})$ simply writes:

$$J_L^{ber}(\mathbf{x}_\tau, \mathbf{y}) = \frac{2}{n}\left(\mathbf{y}^T \log\left(\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}_\tau}\right) + \mathbf{1}^T(\mathbf{H}\mathbf{x}_\tau) - \mathbf{1}^T\mathbf{y}\right) \tag{25}$$

and (24) is still valid. We compare this estimator and the Bardsley estimator, introduced in the paragraph 2.2, on the two synthetic images presented on the figure 4. The algorithm used for this test is the one presented in the next section with the TV regularization. Values given by these two estimators are compared to the optimal value given by the *Mean Square Error* (MSE) defined as:

$$MSE(\mathbf{x}_\tau) = \|\mathbf{x} - \mathbf{x}_\tau\|_2^2 \tag{26}$$

(a) (b) (c)

Figure 4: Synthetic images.

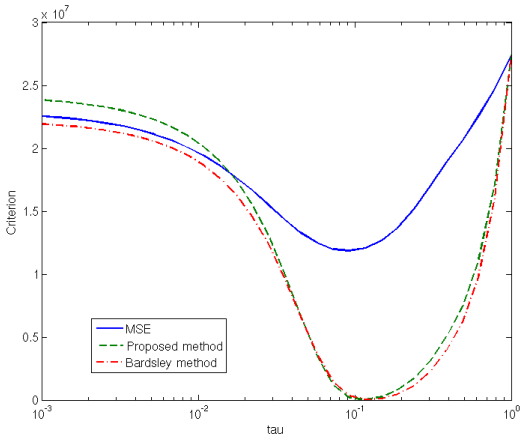


Figure 5: Comparaison of the two methods for the estimation of τ on the synthetic image (a) from the figure 4

where $\mathbf{x}$ is the true image.

The obtained results are shown on the figures 5, 6 and 7. We see that we get a different behavior depending on the test image. On the images (a) and (c), the proposed method (24) gives a value which is close to the value given by the minimization of the MSE. We get $\tau = 0.124$ (target value given by the minimization of the MSE is $\tau = 0.092$) for the image (a) and $\tau = 0.148$ (target value is $\tau = 0.117$) for the image (c). However, we also check that the proposed method may fail on some image. For example on the image (b) we get $\tau = 0.3360$ while the target value is $\tau = 0.0183$. However, the method proposed in [2] also gives a value which is generally too important and tends to oversmooth the image. This behavior has also been noticed in the recent work of [17] with a different estimator tested on a different image. Our experimentations showed that the proposed estimator is generally more efficient on images having a background such as images (a) and (c). We have also observed that the accuracy of the estimator depends on how well the prior is modeled. Clearly images (a) and (c) are well adapted to a Total Variation prior while this prior

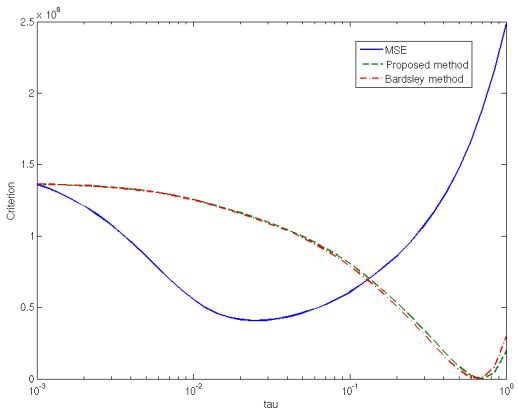


Figure 6: Comparaison of the two methods for the estimation of τ on the synthetic image (b) from the figure 4.

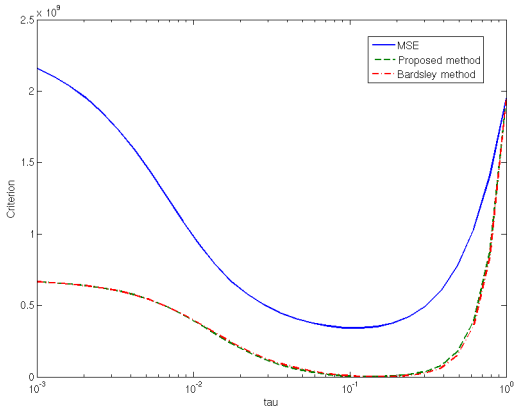


Figure 7: Comparaison of the two methods for the estimation of τ on the synthetic image (c) from the figure 4.

may not suit to image (b). For our experiments, we will use the proposed estimator to select the value of τ.

3.3 Proposed algorithms

By introducing the proposed prior, we want to minimize:

$$J(\mathbf{x}, \mathbf{y}) = J_L(\mathbf{x}, \mathbf{y}) + \tau J_R(\mathbf{x}) = \mathbf{1}^T (\mathbf{H}\mathbf{x}) - \mathbf{y}^T \log[\mathbf{H}\mathbf{x}] + \tau \|\mathbf{W}\mathbf{x}\|_1 \qquad (27)$$

This function is a closed convex function and strongly convex if $\mathbf{y} > 0$ [5]. As often mentioned, this type of problem is not straightforward to solve due to the l^1-norm non-differentiability and due to the presence of operators $\mathbf{H}$ and $\mathbf{W}$.

3.3.1 Richardson-Lucy multiplicative algorithm

The algorithm in its multiplicative form writes (see appendix A):

$$\mathbf{x}_{k+1} = \frac{\mathbf{x}_k}{1 + \tau \mathbf{W}^* \left(\frac{\mathbf{W}\mathbf{x}_k}{|\mathbf{W}\mathbf{x}_k|} \right)} . \left\{ \mathbf{H}^* \left[\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}_k} \right] \right\} \tag{28}$$

where $\mathbf{W}^*$ denotes the adjoint operator of $\mathbf{W}$, obtained using the reconstruction scheme with the time-reversed filters of the decomposition scheme. The parameter τ has to be small enough such that the property of non-negativity can be maintained. However, even for a small value of τ, this algorithm is unstable, as the denominator of (28) may be close to 0. We propose in the next paragraph another algorithm to adress this issue.

3.3.2 Alternating direction method

[16] proposed to use the additive form of the previous algorithm to have a stable behaviour regarding to τ, i.e. to minimize (27) using a gradient descent. On this problem, the step descent of a gradient descent would be of order of $O(\epsilon^2)$, where ϵ is a *small* additive constant which prevents $\mathbf{H}\mathbf{x}$ in the logarithm to be equal to 0. In consequence, the resulting algorithm will be very slow. It can be accelerated using one of the framework of [4, 27], but even with these techniques, an accelerated algorithm on this problem can not give competitive computing time. We propose to use instead an algorithm based on the *Alternating Direction Method* (ADM) [19, 28].

The ADM was initially propose to solve the following problem:

$$\begin{array}{rl} \arg\min & f_1(\mathbf{u}) + f_2(\mathbf{v}) \\ \text{subject to} & A\mathbf{u} + B\mathbf{v} = \mathbf{a} \\ & \mathbf{u} \in \mathbb{R}^n, \mathbf{v} \in \mathbb{R}^m \end{array} \tag{29}$$

where:

- $f_1 : \mathbb{R}^n \to \mathbb{R}$ and $f_2 : \mathbb{R}^m \to \mathbb{R}$ are two closed convex functions.
- $\mathbf{A} \in \mathbb{R}^{l \times n}$ and $\mathbf{B} \in \mathbb{R}^{l \times m}$ are two linear transforms.
- $\mathbf{a} \in \mathbb{R}^l$ is a given vector.

This algorithm is based on the minimization of the augmented Lagragian. Using a Lagrange multiplier $\lambda \in \mathbb{R}^l$ for the linear constraint (29), the augmented Lagrangian writes:

$$\mathcal{L}(\mathbf{u}, \mathbf{v}, \lambda) = f_1(\mathbf{u}) + f_2(\mathbf{v}) + \lambda^T(\mathbf{A}\mathbf{u} + \mathbf{B}\mathbf{v} - \mathbf{a}) + \frac{\beta}{2}\|\mathbf{A}\mathbf{u} + \mathbf{B}\mathbf{v} - \mathbf{a}\|_2^2 \tag{30}$$

β is a parameter which controls the linear constraint [21]. This algorithm consists in minimizing the augmented Lagrangian, in an alternating way, subject to $\mathbf{u}$, $\mathbf{v}$, then to λ. The algorithm writes :

Algorithm 1: ADM to solve (29)

Data: Number of iterations N;
A starting point $\mathbf{u}^0 \in \mathbb{R}^n$;
A starting point $\mathbf{v}^0 \in \mathbb{R}^m$;
A starting point $\lambda^0 \in \mathbb{R}^l$;
Value of the parameters $\gamma > 0$ and $\beta > 0$;
Result: $(\mathbf{u}^N, \mathbf{v}^N)$, an estimated of the solution of (29).
begin
  for k *from* 0 *to* $N-1$ **do**
    Step 1. $\mathbf{u}^{k+1} = \begin{array}{rl} \arg\min & \mathcal{L}(\mathbf{u}, \mathbf{v}^k, \lambda^k) \\ \text{subject to} & \mathbf{u} \in \mathbb{R}^m \end{array}$.
    Step 2. $\mathbf{v}^{k+1} = \begin{array}{rl} \arg\min & \mathcal{L}(\mathbf{u}^{k+1}, \mathbf{v}, \lambda^k) \\ \text{subject to} & \mathbf{v} \in \mathbb{R}^n \end{array}$.
    Step 3. $\lambda^{k+1} = \lambda^k + \beta\gamma(\mathbf{A}\mathbf{u}^{k+1} + \mathbf{B}\mathbf{v}^{k+1} - \mathbf{a})$.
  end
end

γ is a relaxation parameter which has to belongs to $]0, \frac{\sqrt{5}+1}{2}[$ [21] to ensure the convergence of the algorithm. We will set this parameter to be equal to 1 in our experiments. β is the parameter which controls the constraint. The algorithm converges for $\forall \beta > 0$, however the speed of convergence strongly depends on this parameter. If β is small, the convergence of the algortihm will be fast but the linear constraint will take many more iterations to be respected. On the contrary if β is high, then the algorithm will be slow but the linear constraint will be quickly respected. Setting this parameter is actually an open problem and, for our expriments, we will set this parameter equal to 0.1.

We show in the next lines how the ADM algorithm can be used to solve the Poissonian deconvolution problem. We recall that our problem is to find:

$$\begin{array}{rl} \arg\min & \mathbf{1}^T(\mathbf{H}\mathbf{x}) - \mathbf{y}^T \log[\mathbf{H}\mathbf{x}] + \tau\|\mathbf{W}\mathbf{x}\|_1 \\ \text{subject to} & \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \geq \epsilon \end{array} \tag{31}$$

First, we can see that this problem is actually equivalent to:

$$\begin{array}{rl} \arg\min & \mathbf{1}^T\mathbf{w} - \mathbf{y}^T \log[\mathbf{w}] + \tau\|\mathbf{z}\|_1 \\ \text{subject to} & \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \geq \epsilon \\ & \mathbf{w} \in \mathbb{R}^n, \mathbf{w} = \mathbf{H}\mathbf{x} \\ & \mathbf{z} \in \mathbb{R}^m, \mathbf{z} = \mathbf{W}\mathbf{x} \end{array} \tag{32}$$

We set:

$$\mathbf{u} = \begin{pmatrix} \mathbf{x} \\ \mathbf{w} \\ \mathbf{z} \end{pmatrix} \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m, \quad \mathbf{A} = -\mathbf{I}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{I} \\ \mathbf{H} \\ \mathbf{W} \end{bmatrix} \tag{33}$$

$$f_1(\mathbf{u}) = \mathbf{1}^T\mathbf{w} - \mathbf{y}^T \log[\mathbf{w}] + \tau\|\mathbf{z}\|_1 + \chi_C(\mathbf{x}), \quad f_2(\mathbf{v}) = 0 \tag{34}$$

with χ_C being the indicator function on the non-empty convex set C:

$$\chi_C(\mathbf{x}) = \begin{cases} 0 & \text{if} \quad \mathbf{x} \in \mathbb{R}^n, \mathbf{x_i} \geq \epsilon \quad \forall \mathbf{i} \in \mathbf{\Omega} \\ \infty & \text{otherwise} \end{cases} \tag{35}$$

Then the problem (31) can be written as:

$$\begin{array}{rl} \arg\min & f_1(\mathbf{u}) \\ \text{subject to} & -\mathbf{u} + \mathbf{B}\mathbf{v} = 0 \\ & \mathbf{u} \in \mathbb{R}^n, \mathbf{v} \in \mathbb{R}^m \end{array} \tag{36}$$

We see that this formulation completely fits into the framework of the ADM method (29). The first step of the algorithm is to find:

$$\mathbf{u}^{k+1} = \begin{array}{rl} \arg\min & \mathcal{L}(\mathbf{u}, \mathbf{v}^k, \lambda^k) \\ \text{subject to} & \mathbf{u} \in \mathbb{R}^n \end{array} \tag{37}$$

From (36) we can write the augmented Lagrangian as:

$$\mathcal{L}(\mathbf{u}, \mathbf{v}, \lambda) = f_1(\mathbf{u}) + \lambda^T(\mathbf{B}\mathbf{v} - \mathbf{u}) + \frac{\beta}{2}\|\mathbf{B}\mathbf{v} - \mathbf{u}\|_2^2 \tag{38}$$

Then (37) becomes:

$$\begin{aligned} \mathbf{u}^{k+1} &= \begin{array}{rl} \arg\min & f_1(\mathbf{u}) + {\lambda^k}^T(\mathbf{B}\mathbf{v}^k - \mathbf{u}) + \frac{\beta}{2}\|\mathbf{B}\mathbf{v}^k - \mathbf{u}\|_2^2 \\ \text{subject to} & \mathbf{u} \in \mathbb{R}^n \end{array} \\ &= \begin{array}{rl} \arg\min & \frac{1}{\beta} f_1(\mathbf{u}) + \frac{1}{2}\|\mathbf{B}\mathbf{v}^k - \mathbf{u} + \frac{\lambda^k}{\beta}\|_2^2 \\ \text{subject to} & \mathbf{u} \in \mathbb{R}^n \end{array} \\ &= \text{prox}_{\frac{1}{\beta} f_1}\left(\mathbf{B}\mathbf{v}^k + \frac{\lambda^k}{\beta} \right) \end{aligned} \tag{39}$$

where prox is the proximal operator defined by [15]:

$$\text{prox}_{\gamma\Psi}\left(\mathbf{x}_0 \right) = \begin{array}{rl} \arg\min & \gamma\Psi(\mathbf{x}) + \frac{1}{2}\|\mathbf{x}_0 - \mathbf{x}\|_2^2 \\ \text{subject to} & \mathbf{x} \in \mathbb{R}^n \end{array} \tag{40}$$

This proximal operator can be computed in closed-form for some functions Ψ. We give here some examples [15]:

- If $\Psi(x) = \|x\|_1$ then $\text{prox}_{\gamma\Psi}\left(\mathbf{x}_0 \right)$ is the soft-thresholding operator $\text{shrink}_\gamma(\mathbf{x}_0)$ of threshold γ given by:

$$\text{shrink}_\gamma(\mathbf{x}_0) = \text{sign}(\mathbf{x}_0)\max(|\mathbf{x}_0| - \gamma, 0) \tag{41}$$

- If $\Psi(x) = \mathbf{1}^T\mathbf{x} - \mathbf{y}^T\log(\mathbf{x})$ then:

$$\text{prox}_{\gamma\Psi}\left(\mathbf{x}_0 \right) = \frac{1}{2}\left(\mathbf{x}_0 - \gamma + \sqrt{(\mathbf{x}_0 - \gamma)^2 + 4\gamma\mathbf{y}} \right) \tag{42}$$

- If $\Psi(x) = \chi_C(\mathbf{x})$ is the indicator function on a convex set C, then:

$$\text{prox}_{\gamma\Psi}\left(\mathbf{x}_0 \right) = \Pi_C(\mathbf{x}_0) \tag{43}$$

  is the orthogonal projection on this set.

As the proximal operator is componentwise, from (39) we get that for any $\mathbf{u} = \begin{bmatrix} \mathbf{x} \\ \mathbf{w} \\ \mathbf{z} \end{bmatrix} \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$:

$$\text{prox}_{\frac{1}{\beta} f_1}\left(\ \mathbf{u}\ \right) = \begin{pmatrix} \max(\mathbf{x}, \epsilon) \\ \frac{1}{2}\left[\mathbf{w} - \frac{1}{\beta} + \sqrt{\left(\mathbf{w} - \frac{1}{\beta}\right)^2 + \frac{4\mathbf{y}}{\beta}}\right] \\ \text{sign}(\mathbf{z}) \max(|\mathbf{z}| - \frac{\tau}{\beta}, 0) \end{pmatrix} \tag{44}$$

The second step of the algorithm is to find:

$$\begin{aligned} \mathbf{v}^{k+1} = & \quad \arg\min \quad \mathcal{L}(\mathbf{u}^{k+1}, \mathbf{v}, \lambda^k) \\ & \text{subject to} \quad \mathbf{v} \in \mathbb{R}^n \\ = & \quad \arg\min \quad {\lambda^k}^T(\mathbf{B}\mathbf{v} - \mathbf{u}^{k+1}) + \frac{\beta}{2}\|\mathbf{B}\mathbf{v} - \mathbf{u}^{k+1}|_2^2 \\ & \text{subject to} \quad \mathbf{v} \in \mathbb{R}^n \\ = & \quad \arg\min \quad \|\mathbf{B}\mathbf{v} - \mathbf{u}^{k+1} + \frac{\lambda^k}{\beta}\|_2^2 \\ & \text{subject to} \quad \mathbf{v} \in \mathbb{R}^n \end{aligned} \tag{45}$$

Then the solution of (45) can be written as the solution of the following linear system:

$$\mathbf{B}^*\mathbf{B}\mathbf{v}^{k+1} = \mathbf{B}^*\left(\mathbf{u}^{k+1} - \frac{\lambda^k}{\beta} + \mathbf{a}\right) \tag{46}$$

which can always be solved with a conjugate gradient method [5] since $(\mathbf{B}^*\mathbf{B})^* = \mathbf{B}^*\mathbf{B}$. However, using this technique, for each iteration of this inner loop, we have to compute $\mathbf{B}^*\mathbf{B}$. Even if the conjugate gradient loop only need 6 to 7 iterations to give a solution of precision 10^{-5}, this is really costly for a 3D image and leads to high computing time. But (46) can be exactly solved depending on the structure of the matrices $\mathbf{H}$ and $\mathbf{W}$. Indeed, there is a great interest in using a (normalized) tight-frame (i.e. a transform $\mathbf{W}$ such that $\mathbf{W}^*\mathbf{W} = \mathbf{I}$) like the DTCW transform, Curvelets [6], redundant wavelet transform, or both as in [17]. As the convolution matrix $\mathbf{H}$ can, most of the time, be well implemented using the Fast Fourier Transform (FFT), we only consider tight-frame transform and thus (46) can be written as:

$$\mathbf{v}^{k+1} = (\mathbf{H}^*\mathbf{H} + 2\mathbf{I})^{-1}\mathbf{B}^*\left(\mathbf{u}^{k+1} - \frac{\lambda^k}{\beta} + \mathbf{a}\right) \tag{47}$$

which can be easily implemented using FFT. Finally, the proposed algorithm writes:

Algorithm 2: ADM to solve (31)

Data: Number of iterations N;
A starting point $\mathbf{u}^0 \in \mathbb{R}^n$;
A starting point $\mathbf{v}^0 \in \mathbb{R}^m, \mathbf{t}^0 = \mathbf{B}\mathbf{v}^0 \in \mathbb{R}^l$;
A starting point $\lambda^0 \in \mathbb{R}^l$;
Value of the parameters $\gamma > 0$ and $\beta > 0$;
Result: $\mathbf{v}^N$ an estimated of the solution of (31).
begin
  for k *from* 0 *to* $N-1$ **do**
    Step 1. $\mathbf{u}^{k+1} = \text{prox}_{\frac{1}{\beta}f_1}\left(\ \mathbf{t}^k + \frac{\lambda^k}{\beta}\ \right)$
    Step 2. $\mathbf{v}^{k+1} = (\mathbf{H}^*\mathbf{H} + 2\mathbf{I})^{-1}\mathbf{B}^*\left(\mathbf{u}^{k+1} - \frac{\lambda^k}{\beta}\right)$
    Step 3. $\mathbf{t}^{k+1} = \mathbf{B}\mathbf{v}^{k+1}$
    Step 4. $\lambda^{k+1} = \lambda^k + \beta\gamma(\mathbf{t}^{k+1} - \mathbf{u}^{k+1})$
  end
end

The interesting point is that this algorithm converges even if (46) is not solved exactly (to be more specific, this algorithm converges if the errors are summable [19]). In pratice, this algorithm converges to a solution of (31) in 100 iterations. However, the computing time strongly depends on the regularization term. For example, using the TV regularization, this algorithm converges in 25 minutes (for an image with a size of 256*256*64 voxels). With the DTCW regularization (or the Curvelets), the computed time is much longer and about 1h15 for the same number of iterations.
The main drawback of this algorithm, and more generally algorithms which use an augmentation of the size of the problem, is that it needs some memory. This allows, however, to get a reasonable computing time.

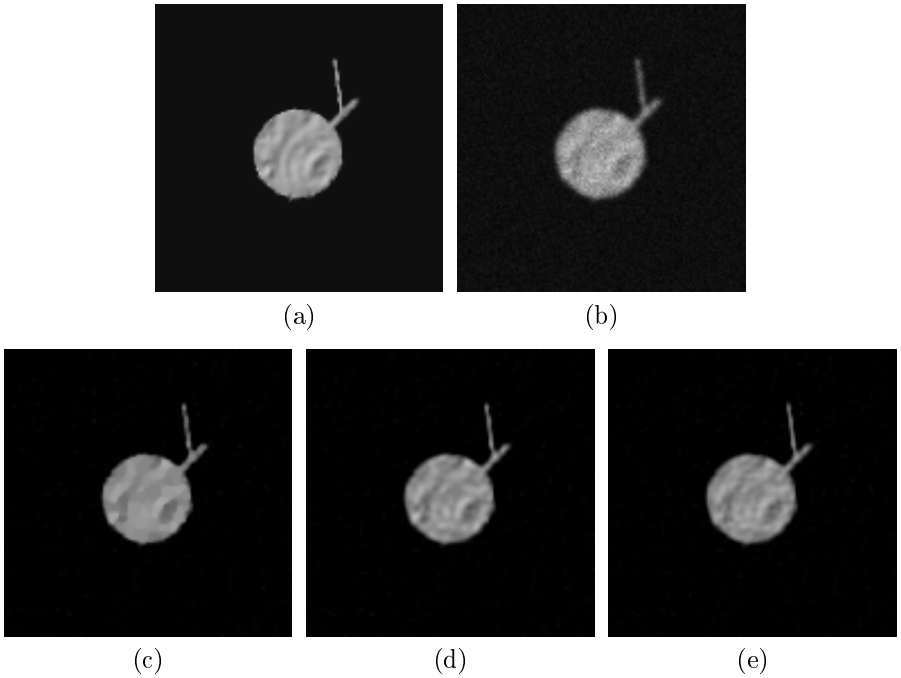


Figure 8: Restoration of a blur and noisy synthetic image. (a) is the original image, (b) is the degraded version ($PSNR = 28.60\ dB$), (c) is the result obtained with the Total Variation prior ($PSNR = 33.87\ dB$), (d) is the result obtained with the proposed prior ($PSNR = 33.65\ dB$), (e) is the result obtained with the curvelets prior ($PSNR = 33.49\ dB$).

4 Results

In this section, we show some results on real and synthetic data using the ADM algorithm with the DTCW regularization. We compare these results to the ADM algorithm with the TV and the Curvelets [6] regularization.

4.1 Results on synthetic data

We first compare the improvement of using wavelets as regularizing operator compared to the TV regularization on the synthetic images (a) and (b) of figure 4. Here, the regularization parameter τ has be chosen such that it minimizes the MSE, in order to evaluate the perfomances of each regularization term. The results obtained on these images are presented on figures 8 and 9.

On these images, wavelets priors allow to retrieve more details than the TV prior. Visually, we can see that the thin elements are better retrieved, in particular we can distinguish the details of the surface of the object. However, as wavelets regularizations, we may have some artifacts and contours may be slightly smoothed. This is quite apparent on the results provided by the Curvelets regularization at the surface of the object (see figure 9). Of course,

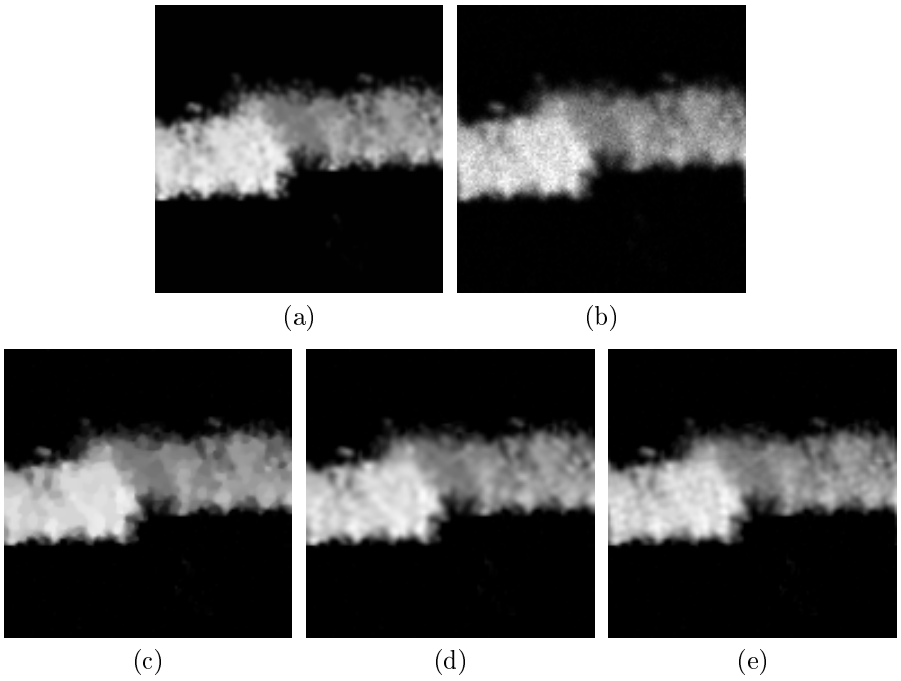


Figure 9: Restoration of a blur and noisy synthetic image. (a) is the original image, (b) is the degraded version ($PSNR = 29.43\ dB$), (c) is the result obtained with the Total Variation prior ($PSNR = 33.83\ dB$), (d) is the result obtained with the proposed prior ($PSNR = 34.53\ dB$), (e) is the result obtained with the curvelets prior ($PSNR = 33.56\ dB$).

Curvelets have been designed to represent curves, and are thus not very efficient to retrieve the texture of the surface of the object, though it is widely used in image restoration.

4.2 Results on real data

We propose to test the algorithm on the restoration of the real images presented on figures 10 and 12. On each test, the regularizing parameter has be chosen using the method proposed in section 3.2. The microscope is a confocal/multi-photon Zeiss Axiovert 200M, with an internal magnification (given by the manufacturer) of 3.3x. The objective is an immersion oil Apochromat[1] 40x for the first image, 63x for the second, with numerical aperture $NA = 1.4$. The oil refractive index is 1.518 (23^o C). The acquisition software is Zeiss LSM 510 Meta.

Figures 10 and 12 show the results obtained with the proposed estimator and the ADM algorithm regularized with the TV, DTCW and Curvelets priors respectively on a sample of mouse intestine and on a bead. First, we see on both figures that the proposed estimator tends to oversmooth the image. This is quite common for this type of estimator, indeed most of the recent proposed estimator for Poisson deconvolution tends to oversmooth the image [2, 17, 42]. On figure 10 we see that the TV regularization smooths the textures of the cells, sticking it together forming a large pattern (right of the image). The DTCW prior allows to retrieve some details of the cells and preserves the space between it, even if the image retrieved is too smoothed. The DTCW prior is quite efficient and gives details of the inside of the cells (zoom on the figure 11). The result obtained with the Curvelets is however not very good as we get many artifacts on the reconstructed image. As previously said, curvelets are not by themself able to well represent surfaces of objects, and for this reason a dictionnary composed of redundant wavelets and curvelets should be used as in [17]. However, this leads to some implementation problems as a 3D redundant wavelet transform need a huge quantity of memory. For this reason, the proposed prior seems to be well adapted to represent these images and does not require important computation ressources.
The image presented on the figure 10 is interesting as it contains many details that can be retrieved using the proposed prior. However, on a smooth object which does not contain many details (as the one presented on the figure 12), the proposed prior does not bring much more information and smoothes the image.

[1] Apochromat is a chromatic aberration correction. From `http://www.wordreference.com`, that is a microscope objective composed by "*a lens, consisting of three or more elements of different types of glass, that is designed to bring light of three colours to the same focal point, thus reducing its chromatic aberration*".

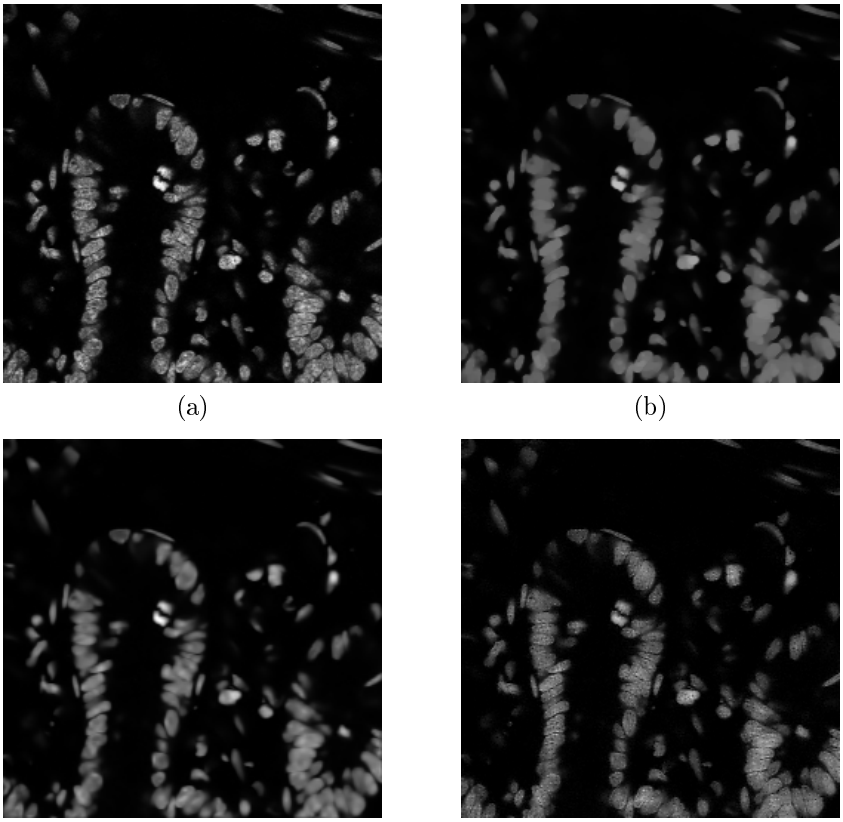
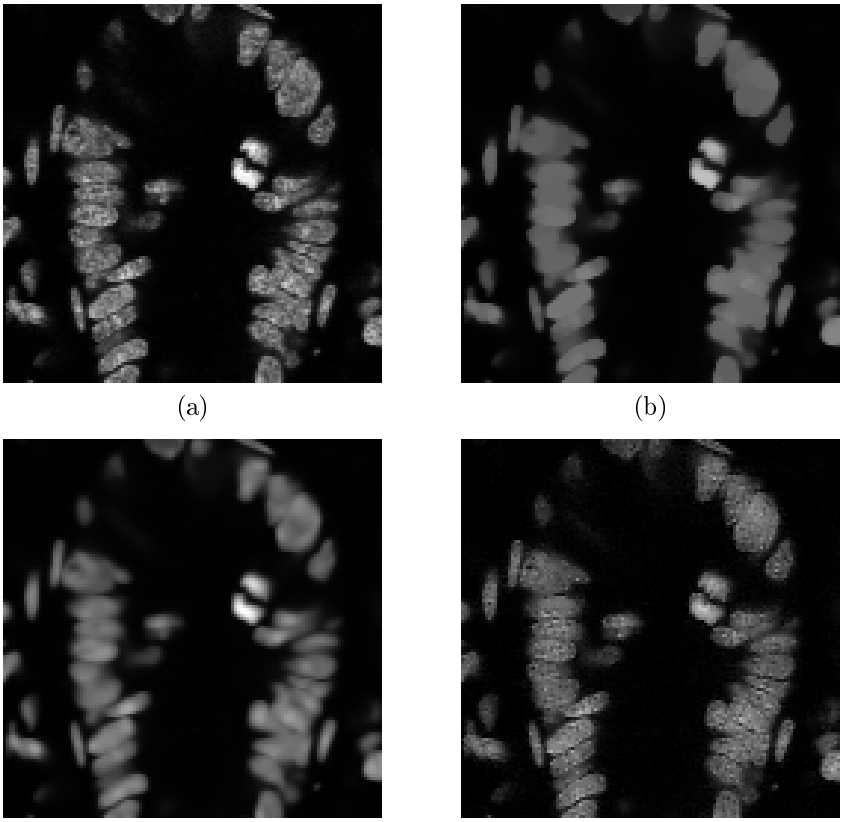


Figure 10: Restoration of a sample of mouse intestin. (a) is the observed image, (b) is the result obtained with the TV prior, (c) is the result obtained with the proposed prior and (d) is the result obtained with curvelets prior.

5 Conclusion

We have proposed a new method for the deconvolution of images corrupted by blur and Poisson noise and applied it on confocal microscopy images. This method includes a wavelet prior which is well adapted to represent the thin structures of the specimens, a method to estimate the value of the regularizing parameter based on a discrepancy principle and a minimization algorithm which is efficient regarding to the important volume of data coming from 3D confocal images. Depending on the structure of the image, the method to select the regularization parameter may give a value which is too important and the image retrieved may be then smoothed too much. Futur works will be focussed on improving the estimator or in developing a constrained formulation of the problem.



(a) (b) (c) (d)

Figure 11: Zoom on the restoration of the sample of mouse intestin. (a) is the observed image, (b) is the result obtained with the TV prior, (c) is the result obtained with the proposed prior and (d) is the result obtained with curvelets prior.

A Richardson-Lucy algorithms

In this appendix, we present the algorithms that are used in this work in detail.

A.1 Non-regularized Richardson-Lucy algorithm

We assume the following image formation model: $\mathbf{y} = \mathcal{P}(\mathbf{Hx})$. $\mathbf{y}$ is the observation, $\mathbf{x}$ is the object, $\mathbf{H}$ stands for the matrix notation of the PSF h of the system and $\mathcal{P}$ is the Poisson distribution. We remind that with a Bayesian approach, we can write:

$$p(\mathbf{x}|\mathbf{y}) = p(\mathbf{y}|\mathbf{x}).\frac{p(\mathbf{x})}{p(\mathbf{y})} \tag{48}$$

where $p(\mathbf{y}|\mathbf{x})$ is the likelihood probability, $p(\mathbf{x}|\mathbf{y})$ the a posteriori probability, and $p(\mathbf{x})$ a prior model on the object. The likelihood probability could be

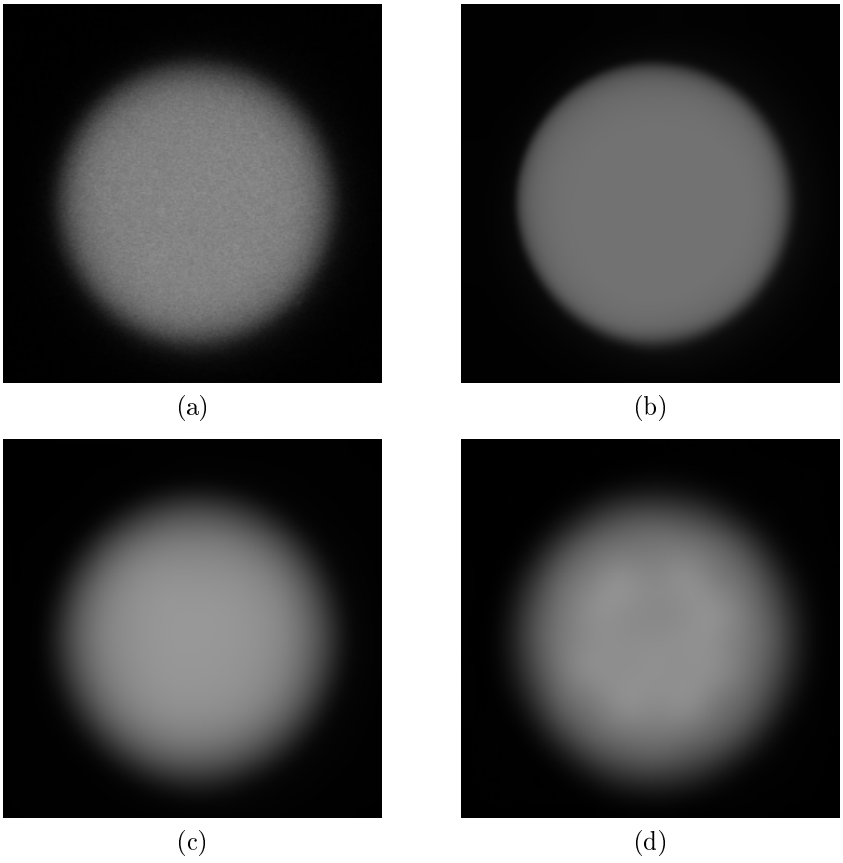


Figure 12: Restoration of a bead. (a) is the original image, (b) is the result obtained with the TV prior, (c) is the result obtained with the proposed prior and (d) is the result obtained with the Curvelets prior.

expressed as:

$$p(\mathbf{y}|\mathbf{x}) = \prod_{\mathbf{i}\in\mathbf{\Omega}} \left(\frac{\left[(\mathbf{Hx})^{\mathbf{y}}\right]_{\mathbf{i}} . \exp\left[-\mathbf{Hx}\right]_{\mathbf{i}}}{\mathbf{y}_{\mathbf{i}}!} \right) \tag{49}$$

One way to solve this problem is to maximize this likelihood probability, which is equivalent to minimize:

$$-\log p(\mathbf{y}|\mathbf{x}) = \mathbf{1}^T\mathbf{Hx} - \mathbf{y}^T \log\left[\mathbf{Hx}\right] + \log\left(\mathbf{y}!\right) \tag{50}$$

The term $\log(\mathbf{y}!)$ in (50) is a constant relatively to $\mathbf{x}$, we can thus define the new functional to minimize as:

$$J_p(\mathbf{x}) = \mathbf{1}^T\mathbf{Hx} - \mathbf{y}^T \log\left[\mathbf{Hx}\right] \tag{51}$$

Considering a small perturbation $\rho\mathbf{s}$ of $\mathbf{x}$ ($\rho > 0$), (51) becomes:

$$J_p(\mathbf{x}+\rho\mathbf{s}) = \mathbf{1}^T\mathbf{H}\left(\mathbf{x}+\rho\mathbf{s}\right) - \mathbf{y}^T \log\left[\mathbf{H}\left(\mathbf{x}+\rho\mathbf{s}\right)\right] \tag{52}$$

Previous equation becomes:

$$\begin{aligned} J_p(\mathbf{x}+\rho\mathbf{s}) &= \mathbf{1}^T\mathbf{H}\mathbf{x} + \rho\mathbf{1}^T\mathbf{H}\mathbf{s} - \mathbf{y}^T\log\left[\mathbf{H}\mathbf{x}\left(1+\rho\frac{\mathbf{H}\mathbf{s}}{\mathbf{H}\mathbf{x}}\right)\right] \\ &\simeq \mathbf{1}^T\mathbf{H}\mathbf{x} + \rho\mathbf{1}^T\mathbf{H}\mathbf{s} - \mathbf{y}^T\log\left[\mathbf{H}\mathbf{x}\right] - \rho\mathbf{y}^T\frac{\mathbf{H}\mathbf{x}}{\mathbf{H}\mathbf{s}} \\ &= J_p(\mathbf{x}) + \rho\left(\mathbf{1}^T\mathbf{H}\mathbf{s} - \mathbf{y}^T\frac{\mathbf{H}\mathbf{x}}{\mathbf{H}\mathbf{s}}\right) \end{aligned} \tag{53}$$

The derivative of J_p can be expressed as:

$$\rho\mathbf{s}^T\frac{\partial J_p}{\partial\mathbf{x}}(\mathbf{x}) = \lim_{\rho\to 0}\frac{J_p\left(\mathbf{x}+\rho\mathbf{s}\right) - J_p(\mathbf{x})}{\rho} \tag{54}$$

Expressing the second term of (53) as a scalar product leads to:

$$\begin{aligned} \mathbf{1}^T\mathbf{H}\mathbf{s} - \mathbf{y}^T\frac{\mathbf{H}\mathbf{s}}{\mathbf{H}\mathbf{x}} &= \mathbf{1}^T\mathbf{H}\mathbf{s} - \left[\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}}\right]^T\mathbf{H}\mathbf{s} \\ &= \mathbf{s}^T\left(\mathbf{H}^*\mathbf{1}\right) - \mathbf{s}^T\left(\mathbf{H}^*\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}}\right) \\ &= \mathbf{s}^T\left(\mathbf{H}^*\mathbf{1} - \mathbf{H}^*\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}}\right) \end{aligned} \tag{55}$$

$\mathbf{H}^*$ denotes the adjoint operator of $\mathbf{H}$. We are now able to express $\nabla_{\mathbf{x}}J_p = \frac{\partial J_p}{\partial\mathbf{x}}$ as:

$$\nabla J_p(\mathbf{x}) = \mathbf{H}^*\left[\mathbf{1} - \frac{\mathbf{y}}{\mathbf{H}\mathbf{x}}\right] \tag{56}$$

We can now minimize (56), i.e. solve $\nabla J_p(\mathbf{x}) = 0$:

$$\mathbf{1}^T\mathbf{H}^*\mathbf{1} - \mathbf{1}^T\mathbf{H}^*\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}} = 0 \tag{57}$$

We assume that the PSF is normalized to 1, then we have:

$$\mathbf{H}^*\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}} = \mathbf{1} \tag{58}$$

By solving (58) in a multiplicative way, we get to the Richardson-Lucy iterative algorithm, by assuming that at convergence, the ratio $\frac{\mathbf{x}_{k+1}}{\mathbf{x}_k}$ is $\mathbf{1}$:

$$\mathbf{x}_{k+1} = \mathbf{x}_k.\left\{\mathbf{H}^*\left[\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}}\right]\right\} \tag{59}$$

A.2 Tikhonov-Miller regularization

Tikhonov-Miller have introduced a regularization term based on the l^2-norm of the image. It could be modelled as an a priori model on the object, thus we can write the statistics on the object as $p(\mathbf{x}) = Ce^{-\tau||\nabla\mathbf{x}||_2^2}$. τ is the regularization parameter for regularization term, and C is the normalization constant for the probability. To take into account the object model, we have to maximize the a posteriori probability $p(\mathbf{x}|\mathbf{y})$. We have to minimize the functional:

$$J(\mathbf{x}) = \mathbf{1}^T\mathbf{H}\mathbf{x} - \mathbf{y}^T\log\left[\mathbf{H}\mathbf{x}\right] + \tau\|\nabla\mathbf{x}\|_2^2 \tag{60}$$

The derivative of the first term gives the same result as(56), and the minimization of the regularization term (that we call J_r) gives:

$$\begin{aligned} J_r(\mathbf{x}+\rho\mathbf{s}) &= \|\nabla(\mathbf{x}+\rho\mathbf{s})\|_2^2 && (61)\\ &\simeq \|\nabla\mathbf{x}\|_2^2 + 2\rho(\nabla\mathbf{s})^T(\nabla\mathbf{x}) && (62)\\ &= J_r(\mathbf{x}) + 2\rho(\nabla\mathbf{s})^T(\nabla\mathbf{x}) && (63)\\ &= J_r(\mathbf{x}) - 2\rho\mathbf{s}^T\mathrm{div}(\nabla\mathbf{x}) && (64)\\ & && (65) \end{aligned}$$

where $\nabla^* = -\mathrm{div}$ is the divergence operator. We can apply for J_r the same equation as (54) to finally obtain the functional minimized relative to o:

$$\mathbf{1}^T\mathbf{H}^*\mathbf{1} - \mathbf{H}^*\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}} - 2\tau\mathbf{1}^T\mathrm{div}(\nabla\mathbf{x}) \tag{66}$$

To solve this equation by using a multiplicative approach (59), we obtain a regularized version of the Richardson-Lucy algorithm:

$$\mathbf{x}_{k+1} = \frac{\mathbf{x}_k}{\mathbf{1} - 2\tau\mathrm{div}(\nabla\mathbf{x}_k)}.\left\{\mathbf{H}^*\left[\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}_k}\right]\right\} \tag{67}$$

A.3 Total Variation regularization

The Total Variation is a regularization method introduced by [34] on the gradient of the image. When used as a prior model for the object, we have to minimize the following functional:

$$\begin{array}{rl} \arg\min & J_p(\mathbf{x}) + \tau\|\nabla\mathbf{x}\|_1 \\ \text{subject to} & \mathbf{x}\in\mathbb{R}^n \end{array} \tag{68}$$

τ being the regularization parameter. We first replace the Total Variation member by its smooth approximation, by adding a *small* constant μ^2 that avoid some problems near the origin. Expanding the Total Variation term $J_r(\mathbf{x}) = \sum_{\mathbf{i}\in\mathbf{\Omega}}\sqrt{|(\nabla\mathbf{x})_\mathbf{i}|^2 + \mu^2}$ of a small perturbation $\mathbf{x}+\rho\mathbf{s}$ of $\mathbf{x}$ yields:

$$
\begin{aligned}
J_r(\mathbf{x}+\rho\mathbf{s}) &= \sum_{\mathbf{i}\in\mathbf{\Omega}} \sqrt{|(\nabla\mathbf{x})_\mathbf{i}|^2 + \rho^2|(\nabla\mathbf{s})_\mathbf{i}|^2 + 2\rho(\nabla\mathbf{x})_\mathbf{i}(\nabla\mathbf{s})_\mathbf{i} + \mu^2} \\
&\simeq \sum_{\mathbf{i}\in\mathbf{\Omega}} \sqrt{|(\nabla\mathbf{x})_\mathbf{i}|^2 + \mu^2}.\sqrt{1 + 2\rho\frac{(\nabla\mathbf{x})_\mathbf{i}(\nabla\mathbf{s})_\mathbf{i}}{|\nabla\mathbf{x})_\mathbf{i}|^2 + \mu^2}} \\
&\simeq J_r(\mathbf{x}) + \rho\sum_{\mathbf{i}\in\mathbf{\Omega}} \frac{(\nabla\mathbf{x})_\mathbf{i}}{\sqrt{|(\nabla\mathbf{x})_\mathbf{i}|^2 + \mu^2}}(\nabla\mathbf{s})_\mathbf{i} \\
&\simeq J_r(\mathbf{x}) + \rho\sum_{\mathbf{i}\in\mathbf{\Omega}} \frac{(\nabla\mathbf{x})_\mathbf{i}}{|(\nabla\mathbf{x})_\mathbf{i}|}(\nabla\mathbf{s})_\mathbf{i} \\
&= J_r(\mathbf{x}) + \rho\left(\frac{\nabla\mathbf{x}}{|\nabla\mathbf{x}|}\right)^T\nabla\mathbf{s} \\
&= J_r(\mathbf{x}) + \rho\left(\nabla^*\frac{\nabla\mathbf{x}}{|\nabla\mathbf{x}|}\right)^T\mathbf{s} \\
&= J_r(\mathbf{x}) - \rho\left[\mathrm{div}\left(\frac{\nabla\mathbf{x}}{|\nabla\mathbf{x}|}\right)\right]^T\mathbf{s}
\end{aligned}
\tag{69}
$$

The minimization of $J_p(\mathbf{x})$ is the same as (56) and gives, for the first terms:

$$
\mathbf{1}^T\mathbf{H}^*\mathbf{1} - \mathbf{H}^*\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}} - \tau\mathrm{div}\left(\frac{\nabla\mathbf{x}}{|\nabla\mathbf{x}|}\right) = 0 \tag{70}
$$

If we derive a multiplicative algorithm in the same way as we obtained (59), we have Richarson-Lucy regularized with Total Variation:

$$
\mathbf{x}_{k+1} = \frac{\mathbf{x}_k}{\mathbf{1} - \tau\mathrm{div}\left(\frac{\nabla\mathbf{x}_k}{|\nabla\mathbf{x}_k|}\right)}.\left\{\mathbf{H}^*\left[\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}_k}\right]\right\} \tag{71}
$$

A.4 Dual-Tree Complex Wavelet regularization

The DTCW is a wavelet transform introduced in [35]. Using it as a prior, the functionnal becomes:

$$
\begin{array}{ll}
\arg\min & J_p(\mathbf{x}) + \tau\|\mathbf{W}\mathbf{x}\|_1 \\
\text{subject to} & \mathbf{x}\in\mathbb{R}^n
\end{array}
\tag{72}
$$

This is the same calculation as the one for the Total Variation. By setting $\mathbf{W} = \nabla$, we get:

$$
\mathbf{1}^T\mathbf{H}^*\mathbf{1} - \mathbf{H}^*\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}} + \tau\mathbf{W}^*\left(\frac{\mathbf{W}\mathbf{x}}{|\mathbf{W}\mathbf{x}|}\right) = 0 \tag{73}
$$

where $\mathbf{W}^*$ denotes the adjoint operator of $\mathbf{W}$. Using a multiplicative algorithm as (59), the Richardson-Lucy algorithm for the Dual-Tree Complex Wavelet prior writes:

$$
\mathbf{x}_{k+1} = \frac{\mathbf{x}_k}{\mathbf{1} + \tau\mathbf{W}^*\left(\frac{\mathbf{W}\mathbf{x}_k}{|\mathbf{W}\mathbf{x}_k|}\right)}.\left\{\mathbf{H}^*\left[\frac{\mathbf{y}}{\mathbf{H}\mathbf{x}_k}\right]\right\} \tag{74}
$$

B Anisotropic Wavelet Transform

As mentioned earlier, images in microscopy don't have the same resolution in the three directions. This may introduce an error in the subband regularization introduced in the section 3.1 as the three direction have a different resolution. In this section, we first recall the dyadic wavelet transform and then propose an anistropic version of the complex wavelet transform. We also show some results on real data using the proposed deconvolution method.

B.1 Discrete Wavelet Transform

The following is mainly recalled from [25]. Notations are simplified to consider real signals. A wavelet decomposition is a representation of a signal $x(t)$ on an orthonormal basis $\psi_{j,k}(t)$ ($(j,k) \in \mathbb{Z} \times \mathbb{Z}$):

$$d_{j,k} = \int_{\mathbb{R}} \psi_{j,k}(t) x(t) dt \tag{75}$$

where $d_j[k]$ are the wavelet coefficients. The signal $x(t)$ can be constructed back from the wavelet coefficients:

$$x(t) = \sum_{k \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} d_j[k] \psi_{j,k}(t) \tag{76}$$

Each basis function $\psi_{j,k}(t)$ defines a subspace $\mathbf{W}_j$ and is built by shifting and stretching the mother wavelet $\psi(t)$:

$$\psi_{j,k}(t) = \sqrt{2^{-j}} \psi\left(2^{-j}t - k\right) \tag{77}$$

Wavelets define a certain level of "detail" of the signal $x(t)$. This level of detail increased as $j \to \infty$ and $\mathbf{W}_j$ gets smaller ($\mathbf{W}_{j+1} \subset \mathbf{W}_j$). A family of scale functions $\phi_{j,k}(t)$ (defining a subspace $\mathbf{V}_j$) can be constructed to get the "approximations" coefficients $c_j[k]$ of the signal $x(t)$:

$$\phi_{j,k}(t) = \sqrt{2^{-j}} \phi\left(2^{-j}t - k\right) \tag{78}$$

Since $\mathbf{V}_1 \subset \mathbf{V}_0$ and $\phi_{0,k}(t) \in \mathbf{V}_0$, we can find some filter $h[k]$ such that $\phi_{1,0}(t)$ is a linear combination of $\phi_{0,k}(t)$:

$$\phi(t) = \sqrt{2} \sum_{k \in \mathbb{Z}} h[k] \phi(2t - k) \tag{79}$$

$\mathbf{W}_{j+1}$ actually complements $\mathbf{V}_{j+1}$ to get to $\mathbf{V}_j$, hence $\mathbf{V}_j = \mathbf{V}_{j+1} \oplus \mathbf{W}_{j+1}$. So, since $\mathbf{W}_1 \subset \mathbf{V}_0$, we can find $g[k]$ such that:

$$\psi(t) = \sqrt{2} \sum_{k \in \mathbb{Z}} g[k] \phi(2t - k) \tag{80}$$

Let $c_0[k]$ be the coefficients at the level 0, we have:

$$x(t) = \sum_k c_0[k] \phi_{0,k}(t) \tag{81}$$

As $\mathbf{V}_0 = \mathbf{V}_1 \oplus \mathbf{W}_1$, this last equation can also be written as:

$$x(t) = \sum_k c_0[k]\phi_{0,k}(t) = \sum_k c_1[k]\phi_{1,k}(t) + \sum_k d_1[k]\psi_{1,k}(t) \tag{82}$$

Considering an orthonormal basis, we can write that:

$$c_1[k] = \langle x(t), \phi_{1,k} \rangle = \sum_n c_0[n]\langle \phi_{0,n}(t), \phi_{1,k}(t) \rangle$$

$$= \sum_n c_0[n] \sum_m h[m]\langle \phi(t-n), \phi(t-m-2k) \rangle = \sum_n c_0[n]h[n-2k] \tag{83}$$

The approximation coefficients at the level 1 can thus be obtained by filtering the approximation coefficients at the level 0 with the time-reversed filter of h followed by a downsampling operation of a factor 2. Same result can be achieved with the detail coefficients:

$$d_1[k] = \sum_n c_0[n]g[n-2k] \tag{84}$$

It is quite common to consider that the approximation coefficients at the level 0 are the discretization of the signal $x(t)$, i.e. $c_0[k] = x[k]$ leading the decomposition scheme presented on the figure 13. This decomposition scheme can easily be extended to 2D or 3D signals sampled at the same rate in each dimension. In the next section, we propose a decomposition scheme when the sampling rate is different in one direction as in confocal microscopy.

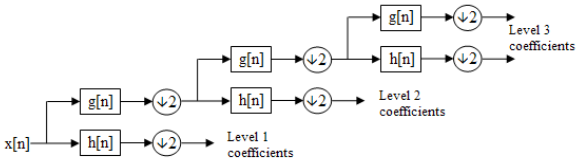


Figure 13: Decomposition scheme of the Discret Wavelet Transform.

B.2 Problem statement

We consider a 3D image $i(x, y, z)$ sampled at rate Δ_x, Δ_y and Δ_z. Our observation is thus the 3D matrix $I(m, n, o) = i(m\Delta_x, n\Delta_y, o\Delta_z)$ with $(n, m, o) \in \mathbb{N}^3$. Usually in microscopy, we have $\Delta_x = \Delta_y$ and we can write that $\Delta_z = \alpha\Delta_x$ with $\alpha > 1$. In the proposed deconvolution method, we use our observation to initialize the wavelet transform, i.e. we project the image on the mother scale function (level 0). We consider that:

$$c_0(n, m, o) = I(m, n, o) \tag{85}$$

But we have (considering separable wavelet, i.e. $\phi(x,y,z) = \phi(x)\phi(y)\phi(z)$):

$$\begin{aligned} c_0(n,m,o) &= \iiint_{\mathbb{R}^3} i(x,y,z')\phi(x-m,y-n,z'-o)\,dxdydz' \\ &= \iiint_{\mathbb{R}^3} i(x,y,z')\phi(x-m)\phi(y-n)\phi(z'-o)\,dxdydz' \\ &= i(m\Delta_x, n\Delta_y, o\Delta'_z) \end{aligned} \tag{86}$$

This implies that $\Delta'_z = \Delta_x = \Delta_y$, meaning that the initialization of the wavelet transform should be equal to the image $i(x,y,z)$ sampled at the same rate $(\Delta_x, \Delta_y, \Delta'_z)$ in the three directions. However as our observation is sampled at the rate $(\Delta_x, \Delta_y, \Delta_z)$, we can write:

$$\begin{aligned} I(m,n,o) &= i(m\Delta_x, n\Delta_y, o\alpha\Delta'_z) \\ &= \iiint_{\mathbb{R}^3} i(x,y,\alpha z')\phi(x-m)\phi(y-n)\phi(z'-o)\,dxdydz' \\ &= 2^J \iiint_{\mathbb{R}^3} i(x,y,z)\phi(x-m)\phi(y-n)\phi\left(\frac{z}{2^J}-o\right)\,dxdydz \\ &= \left(\sqrt{2^J}\right)^3 \iiint_{\mathbb{R}^3} i(x,y,z)\phi_{0,m}(x)\phi_{0,n}(y)\phi_{J,o}(z)\,dxdydz \end{aligned} \tag{87}$$

where $J = \log_2(\alpha)$. For the X and Y directions the projection is really made on the mother scale function. However for the Z direction, the initialization represents a projection on a child scale function corresponding to the *rational* J level. In consequence, our observation can be seen as the low-pass coefficients of a wavelet decomposition on J levels only in the Z direction. With an isotropic wavelet transform, the coefficients in Z are thus always at a greater resolution. A subband regularization will then be greater than the optimal value and this may result in smoothing the edges too much. Getting to the same resolution is thus crucial for deconvolution.

It is not possible to project back on the mother scale function as we do not have the high-pass coefficients of the scale J. Setting this coefficients to 0 and reconstruct the image will not add more information as this operation is equivalent to resample the image.

B.3 Proposed solution

B.3.1 Algorithm

In order to get the same resolution, we propose the following method:

- Choose a number of decomposition levels L.
- Find the closest (coarser) integer *dyadic* scale J'.
- Project the image on J' only in the Z direction.
- Compute J' levels of the 2D dyadic wavelet transform on X and Y directions.
- Compute $L - J'$ levels of the 3D dyadic wavelet transform (at this step, the image is sampled at the same rate in the 3 directions).

This method allows to choose a subband-dependent regularization paramater as the 3D directions get to the same resolution. J' is the closest coarser scale, i.e.:

$$J' = \min\{j \in \mathbb{Z} | j \geq J\} \tag{88}$$

The projection on this scale J' is not however obvious.

B.3.2 Making the projection

We consider that the wavelet is separable: $\phi(x, y, z) = \phi(x)\phi(y)\phi(z)$. As $\mathbf{V}_J \subset \mathbf{V}_{J'}$, we can find a filter $r(n)$ such that:

$$\phi_{J,0}(z) = \sum_{n \in \mathbb{Z}} r(n)\phi_{J',n}(z) \tag{89}$$

Then, the approximation coefficients at the scale J' can be expressed as:

$$c_{J'}(.,.,o) = \langle i(.,.,z), \phi_{J',o}(z) \rangle = \sum_{k \in \mathbb{Z}} c_J(.,.,k) \langle \phi_{J,k}(z), \phi_{J',o}(z) \rangle$$

$$= \sum_{k \in \mathbb{Z}} c_J(.,.,k) r(o - 2^{J-J'}k) \tag{90}$$

Thus, the approximation coefficients at the scale J' are obtained by filtering our observation by the filter r, and then subsample by the *rational* factor $2^{J-J'}$. We can also obtain the details coefficients at the scale J' using the same method:

$$d_{J'}(.,.,o) = \sum_{k \in \mathbb{Z}} c_J(.,.,k) s(o - 2^{J-J'}k) \tag{91}$$

where:

$$\psi_{J,0}(z) = \sum_{n \in \mathbb{Z}} s(n)\phi_{J',n}(z) \tag{92}$$

We can see that both approximation and detail coefficients have to be subsampled by the same factor $2^{J-J'}$ which $\in [1, 2]$. Let us introduce integers M and P ($P < M$) such that $\frac{M}{P} = 2^{J-J'}$. Then the decomposition scheme is the one presentend on the figure 14.

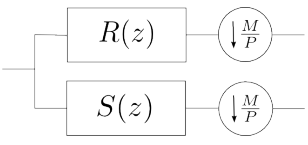
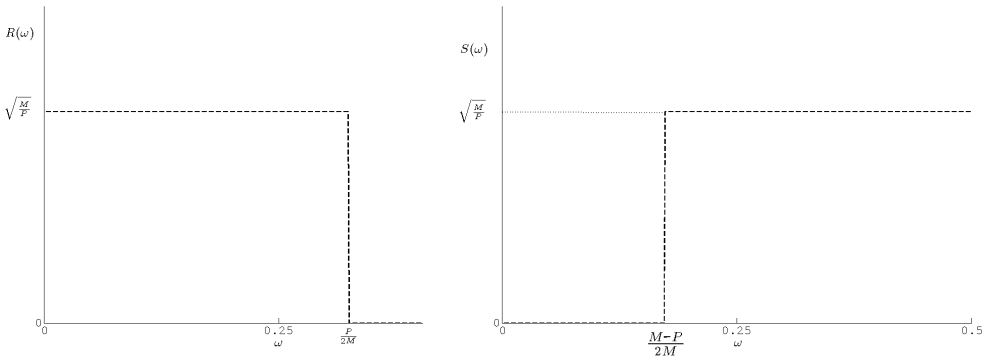


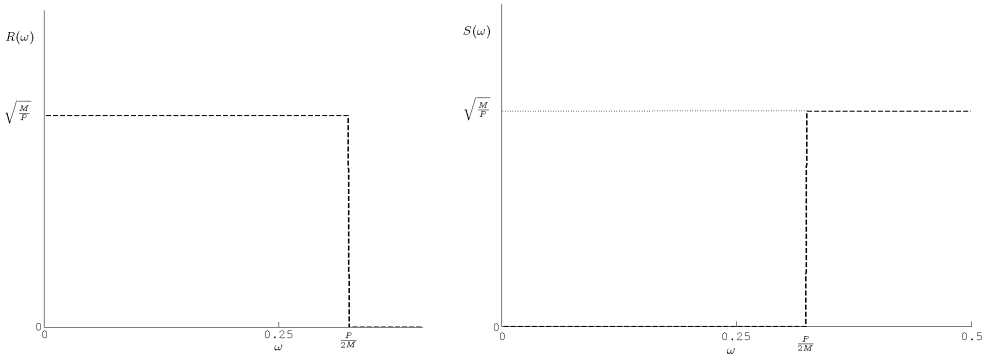
Figure 14: Scheme of the projection.

The low-pass filter r has a cut-frequency $\frac{P}{2M}$ which $\in \left[\frac{1}{4}, \frac{1}{2}\right]$ and the high-pass filter s has a cut-frequency $\frac{M-P}{2M}$ which $\in \left[0, \frac{1}{4}\right]$ (illustrated on the figure 15). These filters actually overlap making the perfect reconstruction property impossible. As the decomposition scheme is an iteration on the low-pass coefficients, we can design the filter s as we want since it allows the perfect reconstruction property. Then, the filters allowing this property are shown on the figure 16.



(a) (b)

Figure 15: Ideal filters r (a) and s (b) in the Fourier domain.



(a) (b)

Figure 16: Ideal filters r (a) and s (b), in the Fourier domain, allowing perfect reconstruction.

In order to get fully decimated, the high-pass branch has to be subsampled by a factor $\frac{M}{M-P}$. However, with this subsample factor, perfect reconstruction can be obtained only if $M = P+1$ [3]. As 3D images in confocal microscopy may be sampled with a large rank of different sampling values, we can barely hold this rule. Thus, the high-pass branch of our projection scheme will be overcomplete and the subsample factor (called Q) of this branch will be chosen such that this redundancy is minimized (while it needs to verify $\frac{P}{M} + \frac{1}{Q} > 1$ to achieve perfect reconstruction). Finally, the projection scheme used in our method is the one presented on the figure 17.
This decomposition scheme has recently been proposed in [3] as a rational (for rational subsample factor) wavelet transform. Thus, we will make a one level decomposition of the wavelet transform proposed in [3] to realize our projection.

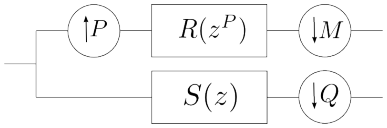


Figure 17: Final scheme for the projection.

B.3.3 Results

We applied this wavelet transform to the restoration of the blur and noisy (Poisson noise) image (b) on the figure 18. Here, the rows of the original image (a) have been subsampled by a factor 2. Under this configuration ($\alpha = 2$), the anisotropic wavelet transform is equal to a 2D wavelet transform where the columns at the first level are not filtered and decimated (the projection is not required here). The results shown on the figure 18 have been obtained using the algorithm proposed in this report. The regularization parameter has, however, be tuned such that it minimizes the mean square error.
The result obtained with the anisotropic wavelet transform contains a slightly more details but also more noise. At the first level of the proposed anistropic wavelet transform, only one direction of the image is filtered and decimated, the other direction thus still contains some noise. Clearly, in order to remove the noise, we should avoid to shut off a whole dimension in the decomposition process. We have however some beliefs that this method is a good start as it shows that taking into account the sampling of the object may bring more details in the restoration of the image.

Figure 18: Results of the anistropic wavelet transform. (a) is the original image with a sampling factor $\alpha = 2$, (b) is the blur and noisy observation, (c) is the result obtained with an isotropic wavelet prior (actually the dual-tree complex wavelet prior that we have proposed in this paper) ($PSNR = 30.7\ dB$) and (d) is the result obtained with an anistropic version of this wavelet transform ($PSNR = 30.2\ dB$).

References

[1] F.J. Anscombe. The transformation, of Poisson, binomial and negative-binomial data. *Biometrika*, 35(3/4):246–254, Dec. 1948.

[2] J. M. Bardsley and J. John Goldes. Regularization parameter selection methods for ill-posed poisson maximum likelihood estimation. *Inverse Problems*, 25(9), 2009.

[3] I. Bayram and I. W. Selesnick. Overcomplete discrete wavelet transforms with rational dilation factors. *IEEE Transactions on Signal Processing*, 57(1):131–145, 2009.

[4] A. Beck and Teboulle M. A fast iterative shrinkage thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1), 2009.

[5] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Mar. 2004.

[6] E. Candès, L. Demanet, D. Donoho, and L. Ying. Fast discrete curvelet transforms. *Multiscale Modeling & Simulation*, 5(3):861–899, 2006.

[7] M. Carlavan, P. Weiss, and L. Blanc-Féraud. Régularité et parcimonie pour les problèmes inverses en imagerie : algorithmes et comparaisons. *Traitement du Signal*, 2010. To appear.

[8] M. Carlavan, P. Weiss, L. Blanc-Féraud, and J. Zerubia. Complex wavelet regularization for solving inverse problems in remote sensing. In *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Cape Town, South Africa, Jul. 2009.

[9] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.*, 20(1-2):89–97, 2004.

[10] C. Chaux, L. Blanc-Féraud, and J. Zerubia. Wavelet-based restoration methods: application to 3d confocal microscopy images. In *Proc. SPIE Conference on Wavelets*, 2007.

[11] C. Chaux, P. L. Combettes, J.-C. Pesquet, and V. R. Wajs. A forward-backward algorithm for image restoration with sparse representations. In *Signal Processing with Adaptative Sparse Structured Representations (SPARS'05)*, pages 49–52, Rennes, France, Nov. 2005.

[12] C. Chaux, L. Duval, and J.-C. Pesquet. Image analysis using a dual-tree m-band wavelet transform. *Image Processing, IEEE Transactions on*, 15(8):2397–2412, Aug. 2006.

[13] C. Chaux, J.-C. Pesquet, and N. Pustelnik. Nested iterative algorithms for convex constrained image recovery problems. *SIAM Journal on Imaging Sciences*, 2(2):730–762, Jun. 2009.

[14] P. L. Combettes and J.-C. Pesquet. A proximal decomposition method for solving convex variational inverse problems. *Inverse Problems*, 24(6), 2008.

[15] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling & Simulation*, 4(4):1168–1200, 2005.

[16] 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.

[17] F.-X. Dupé, J. Fadili, and J.-L. Starck. A proximal iteration for deconvolving poisson noisy images using sparse representations. *IEEE Transactions on Image Processing*, 18(2):310–321, Feb. 2009.

[18] M. Elad, P. Milanfar, and R. Rubinstein. Analysis versus synthesis in signal priors. *Inverse Problems*, 23(3):947–968, 2007.

[19] E. Esser. Applications of lagrangian-based alternating direction methods and connections to split bregman. *UCLA Cam Report*, 2009.

[20] M. A. T. Figueiredo and J. M. Bioucas-Dias. Restoration of poissonian images using alternating direction optimization. *Submitted to IEEE Transactions on Image Processing*, Jan. 2010.

[21] R. Glowinski. *Numerical Methods for Nonlinear variation al Problems*. Springer-Verlag, 1984.

[22] G. H. Golub, M. Heath, and G. Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2):215–223, 1979.

[23] A. Grinvald and I. Z. Steinberg. On the analysis of fluorescence decay kinetics by the method of least-squares. *Analytical Biochemistry*, 59(2):583–598, 1974.

[24] L. B. Lucy. An iterative technique for rectification of observed distributions. *The Astronomical Journal*, 79(6):745–765, 1974.

[25] S. Mallat. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, 2008.

[26] M. Minsky. Memoir on inventing the confocal scanning microscope. *Scanning*, 10:128–138, 1988.

[27] Y. Nesterov. Gradient methods for minimizing composite objective function. *CORE Discussion Paper 2007/76*, 2007.

[28] M. Ng, P. Weiss, and X. M. Yuan. Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods. *ICM Research Report*, Oct. 2009.

[29] G. Pons Bernad, L. Blanc-Féraud, and J. Zerubia. A restoration method for confocal microscopy using complex wavelet transform. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Philadelphia, Pennsylvania, USA, Mar. 2005.

[30] N. Pustelnik, C. Chaux, and J.-C. Pesquet. Hybrid regularization for data restoration in the presence of Poisson noise. In *17th European Signal Processing Conference (EUSIPCO'09)*, Aug. 2009.

[31] S. Ramani, T. Blu, and M. Unser. Monte-Carlo SURE: A Black-Box Optimization of Regularization Parameters for General Denoising Algorithms. *IEEE Transactions on Image Processing*, 17(9):1540–1554, 2008.

[32] S. Ramani, C. Vonesch, and M. Unser. Deconvolution of 3d fluorescence micrographs with automatic risk minimization. In *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pages 732–735, May 2008.

[33] W. H. Richardson. Bayesian-based iterative method of image restoration. *Journal of Optical Society of America*, 62:55–59, 1972.

[34] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[35] I. W. Selesnick, R.G. Baraniuk, and N. G. Kingsbury. The dual-tree complex wavelet transform. *IEEE Signal Processing Magazine*, 22(6):123–151, Nov 2005.

[36] I. W. Selesnick and M. A. T. Figueiredo. Signal restoration with overcomplete wavelet transforms: comparison of analysis and synthesis priors. volume 7446. SPIE, 2009.

[37] S. Setzer, G. Steidl, and T. Teuber. Deblurring poissonian images by split bregman techniques. *Journal of Visual Communication and Image Representation*, 21:193–199, 2010.

[38] C. J. R. Sheppard and C. J. Cogswell. Three-Dimensional Image Formation in Confocal Microscopy. *Journal of of Microscopy*, 159:179–194, Aug. 1990.

[39] G. M. P. van Kempen. *Image Restoration in Fluorescence Microscopy.* PhD thesis, Technische Universiteit Delft - Holland, Jan. 1999.

[40] G. M. P. van Kempen and L. J. Van Vliet. The influence of the regularization parameter and the first estimate on the performance of tikhonov regularized non-linear image restoration algorithms. *Journal of Microscopy*, 13:63–75, Apr. 2000.

[41] G. M. P. van Kempen, L. J. Van Vliet, P. J. Verveer, and H. T. M. Van Der Voort. A quantitative comparison of image restoration methods for confocal microscopy. *Journal of Microscopy*, 185:354–365, Mar. 1997.

[42] R. Zanella, P. Boccacci, L. Zanni, and M. Bertero. Efficient gradient projection methods for edge-preserving removal of poisson noise. *Inverse Problems*, 25(4), 2009.

Centre de recherche INRIA Sophia Antipolis – Méditerranée
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399