3D Microscopy Deconvolution using Richardson-Lucy Algorithm with Total Variation Regularization

Nicolas Dey, Laure Blanc-Féraud, Christophe Zimmer, Pascal Roux, Zvi Kam, Jean-Christophe Olivo-Marin, Josiane Zerubia

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

3D Microscopy Deconvolution using Richardson-Lucy Algorithm with Total Variation Regularization

Nicolas DEY — Laure BLANC-FÉRAUD — Christophe ZIMMER — Pascal ROUX — Zvi KAM — Jean-Christophe OLIVO-MARIN — Josiane ZERUBIA

N° 5272

Juillet 2004

Thème COG

ISSN 0249-6399

INRIA
SOPHIA ANTIPOLIS

3D Microscopy Deconvolution using Richardson-Lucy Algorithm with Total Variation Regularization

Nicolas DEY* , Laure BLANC-FÉRAUD* , Christophe ZIMMER†,
Pascal ROUX‡, Zvi KAM§, Jean-Christophe OLIVO-MARIN† ,
Josiane ZERUBIA*

Thème COG — Systèmes cognitifs
Projet Ariana

Rapport de recherche n° 5272 — Juillet 2004 — 71 pages

Abstract: Confocal laser scanning microscopy is a powerful and increasingly popular technique for 3D imaging of biological specimens. However the acquired images are degraded by blur from out-of-focus light and Poisson noise due to photon-limited detection. Several deconvolution methods have been proposed to reduce these degradations, including the Richardson-Lucy iterative algorithm, which computes a maximum likelihood estimation adapted to Poisson statistics. However this algorithm does not necessarily converge to a suitable solution, as it tends to amplify noise. If it is used with a regularizing constraint (some prior knowledge on the data), Richardson-Lucy regularized with a well-chosen constraint, always converges to a suitable solution. Here, we propose to combine the Richardson-Lucy algorithm with a regularizing constraint based on Total Variation, whose smoothing avoids oscillations while preserving object edges. We show on simulated and real images that this constraint improves the deconvolution results both visually and using quantitative measures. We compare several well-known deconvolution methods to the proposed method, such as standard Richardson-Lucy (no regularization), Richardson-Lucy with Tikhonov-Miller regularization, and an additive gradient-based algorithm.

Key-words: confocal microscopy, 3D image processing, deconvolution, point spread function, multiplicative noise, total variation, ...

* Ariana Group, INRIA/I3S, 2004 route des Lucioles - BP93, 06902 Sophia Antipolis, France
† Quantitative Image Analysis Group, Institut Pasteur, 25-28 rue du Dr. Roux, 75015 Paris, France
‡ Dynamic Imagery Platform Group, Institut Pasteur, 25-28 rue du Dr. Roux, 75015 Paris, France
§ Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel 76100

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 —Télécopie : +33 4 92 38 77 65

Déconvolution en microscopie tridimensionnelle utilisant l'algorithme de Richardson-Lucy régularisé avec la variation totale

Résumé : La microscopie confocale (*Confocal laser scanning microscopy* ou microscopie confocale à balayage laser) est une méthode puissante de plus en plus populaire pour l'imagerie 3D de spécimens biologiques. Malheureusement, les images acquises sont dégradées non seulement par du flou dû à la lumière provenant de zones du spécimen non focalisées, mais aussi par un bruit de Poisson dû à la détection, qui se fait à faible flux de photons. Plusieurs méthodes de déconvolution ont été proposées pour réduire ces dégradations, avec en particulier l'algorithme itératif de Richardson-Lucy, qui calcule un maximum de vraisemblance adapté à une statistique poissonienne. Mais cet algorithme utilisé comme tel ne converge pas nécessairement vers une solution adaptée, car il tend à amplifier le bruit. Si par contre on l'utilise avec une contrainte de régularisation (connaissance a priori sur l'objet que l'on cherche à restaurer, par exemple), Richardson-Lucy régularisé converge toujours vers une solution adaptée, sans amplification du bruit. Nous proposons ici de combiner l'algorithme de Richardson-Lucy avec une contrainte de régularisation basée sur la Variation Totale, dont l'effet d'adoucissement permet d'éviter les oscillations d'intensité tout en préservant les bords des objets. Nous montrons sur des images synthétiques et sur des images réelles que cette contrainte de régularisation améliore les résultats de la déconvolution à la fois qualitativement et quantitativement. Nous comparons plusieurs méthodes de déconvolution bien connues à la méthode que nous proposons, comme Richardson-Lucy standard (pas de régularisation), Richardson-Lucy régularisé avec Tikhonov-Miller, et un algorithme basé sur la descente de gradients (sous l'hypothèse d'un bruit additif gaussien).

Mots-clés : microscopie confocale, traitement d'images 3D, déconvolution, réponse impulsionnelle, bruit multiplicatif, variation totale, ...

Contents

Acknowledgments

This work was funded by the INRIA DeMiTri Cooperative Research Action[1]. The authors would like to thank Charles Kervrann (INRA Jouy-en-Josas / IRISA) and Alain Trubuil (INRA Jouy-en-Josas) for their help; SVI (Scientific Volume Imaging, the Netherlands) for their expertise on the deconvolution problem and for the Huygens software workshop; Jean-françois Aujol for discussions on inverse problems. 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 the images presented in Fig. 21 and 22 of Appendix E.

[1] See `http://www-sop.inria.fr/ariana/demitri.html` for more informations.

Notations and abreviations

$p(.\|.)$	any conditional probability.
$\overline{\zeta}$	the conjugate complex of any complex number ζ.
$\mathcal{Z}^*$	the adjoint operator of any operator $\mathcal{Z}$.
x	a vector of continuous coordinates. Could be (X, Y) in 2D or (X, Y, Z) in 3D.
X, Y, Z	the continuous coordinates in the three dimentions in the direct space.
U, V, W	the continuous coordinates in the three dimentions in the frequency space.
$\mathbf{x}$	the vector of discrete coordinates (in 2D or in 3D).
h	the point spread function (PSF) of the system.
i	the observed image.
o	the initial object.
b	the background.
$\wp(.)$	Poisson distribution.
k	the index for the algorithm iterations.
OTF	optical transfert function, the Fourier transform of the PSF.
PSF	point spread function, the 3D impulse response of the optical system.
TV	Total Variation regularization.
TM	Tikhonov-Miller regularization.
RL	Richardson-Lucy algorithm.
NA	numerical aperture of the objective.

1 Introduction

Confocal microscopy was invented by Marvin Minsky [36]. The first working prototype was constructed in 1955. Confocal microscopy offers several advantages over conventional optical microscopy with its small depth-of-field[2], its reduction of out-of-focus blur, and its full three-dimensional (3D) image scanning ability. For biomedical applications, it can also acquire images of living (fixed or moving) cells, usually labeled with one or more fluorescent probes.

1.1 Confocal microscopy and the need for deconvolution

The confocal laser scanning microscope (CLSM) is an optical fluorescence microscope associated to a laser that scans the specimen in 3D and uses a pinhole to reject most out-of-focus light. The ability of CLSM to image optical sections of thick specimens explains its rapidly increasing use in biological research [38].

Despite the advantages of the CLSM, the quality of confocal microscopy images suffers from two basic physical limitations. First, out-of-focus blur due to the diffraction-limited nature of optical microscopy remains substantial, even though it is reduced compared to widefield microscopy. Second, the confocal pinhole drastically reduces the amount of light detected by the photomultiplier, leading to Poisson noise [38]. The images produced by CLSM can therefore benefit from postprocessing by deconvolution methods designed to reduce blur and/or noise.

1.2 Proposed approach

The aim of this study is to propose a new approach for deconvolution of 3D confocal images. Many deconvolution methods have already been proposed for confocal microscopy, such as Tikhonov-Miller inverse filter [67], the Carrington [67] and Richardson-Lucy (RL) algorithms [31, 44]. The latter has been used extensively in astrophysical or microscopic imaging [67], and is of particular interest for confocal microscopy because it is adapted to Poisson noise. An important drawback of RL deconvolution, however, is that it amplifies noise after a few iterations. This sensitivity to noise can be avoided with the help of regularization constraints, leading to much improved results. Conchello et al. [8, 66] and van Kempen et al. [65] have presented a RL algorithm using energy-based regularization applied to biological images. Conchello's regularization term leaves oscillations introduced by RL iterations in homogeneous areas. Tikhonov-Miller based term, on the contrary, regularizes too much, resulting in smoothed edges.

We propose the following approach to our problem: we first define an image formation model, and the statistics describing the image; we then propose a model for the point spread function (PSF), and a deconvolution algorithm based on RL with Total Variation (TV) regularization. This approach allows to avoid the noise amplification of the non-regularized

[2] Physical terms are defined in Appendix A.

RL method, while preserving the edges of the image. The validation step is performed by testing this method on simulated and real data, and comparing it to the results given by well-known deconvolution algorithms: standard RL and RL with Tikhonov-Miller regularization (see Appendix C).

Which is new in the method proposed in this research report is the regularization constraint: we use the Total Variation semi-norm, which helps to avoid the noise amplification during the deconvolution. Compared to other kinds of regularization, the main advantages of the Total Variation are that it preserves the edges in the image, and smoothes homogeneous areas.

1.3 Contents

In section 2, we present different pieces of work closely related to our subject. We review some of the existing deconvolution methods in confocal microscopy. In section 3, we first describe the image formation in confocal microscopy, and we also present the image formation model we are using. We insist on the image statistics (Poisson process) and the point spread function modelling. Then we propose the theory of the deconvolution method we are using in this paper. It is based on Richardson-Lucy iterative algorithm and is regularized with Total Variation. We also present the quantitative criteria that are used to compare the results on simulated and real data. We show results on simulated data in section 4, and results on real data in section 5. We conclude in section 6, discussing the advantages and the limitations of the method. We present some further tracks that should be investigated to enhance the results or fully validate the method.

2 State of the art

In this section, we present different pieces of work that are very close to or related to this paper. We first present some denoising techniques in confocal microscopy (section 2.1). We then present some methods to model or measure the point spread functions (PSF) (section 2.2). A good PSF estimation is essential for any non-blind deconvolution algorithm presented in section 2.3. We finally present some works about the Total Variation (section 2.4), the regularization term we have chosen to introduce in our model.

2.1 Denoising in confocal microscopy

As a first approximation, 3D confocal images are not blurred compared to other microscopy type (see Appendix A): the resolution of such microscopes is very fine [5]. If the blur is not important relatively to the noise [39], one can only restore the images using denoising methods without deblurring. For instance, if the resolution of the image is approximately equal or less than the size of the PSF, the effects of the blur are important and noticeable. If the resolution is too low compared to the PSF size, the effects of blur will be negligible and only a denoising will be necessary.

In [67], van Kempen et al. apply a Gaussian filter as a pre-processing to their raw data. After this, they use a deconvolution method to restore their data. They show that the results are improved compared to deconvolved images with no pre-filtering.

Boutet de Monvel et al. [12] propose a denoising method for confocal image stacks, using Daubechies' wavelets for each direction, before applying a MAP or a non-regularized Richardson-Lucy method for deconvolution. Wavelet denoising offers an alternative method of regularization for deconvolution. Wavelet denoising followed by RL deconvolution provides a more efficient and more adaptive denoising, because denoising could be performed in one time, by wavelet transform, and then noise-dependant coefficient removal, before an inverse wavelet transform.

In [71], Willett et al. use a multiscale approach based on platelets to denoise 2D images in the presence of a Poisson noise. Platelets are localized functions at various scale, locations, and orientations [70], which produce piecewise linear image approximations. This platelet-based method is very well suited to Poisson noise and preserves edges in the images. They apply their method to several kinds of images with Poisson noise, particularly on phantom and real 2D confocal images. Their method is applied to phantom and real data denoising, and it gives better results than those obtained with Haar wavelet or wedgelet denoising algorithms. They also use platelet-based method for deblurring 2D phantom confocal images [70, 71].

In [28], Kervann and Trubuil propose a method to denoise confocal images as a pre-processing before using a deconvolution algorithm based on Richardson-Lucy. In their paper, they only present the denoising technique. It is based on a locally piecewise constant modelling of the image with an adaptive choice of window around each pixel. They model the image distribution as a Poisson distribution i, and apply the Anscombe trans-

form $I_{Ans}(i,j,k) = 2\sqrt{i(i,j,k) + \frac{3}{8}}$ to obtain a Gaussian distribution I_{Ans} with variance $\sigma^2 = 1$. This way, Kervann and Trubuil are able to use well-known methods in presence of Gaussian noise, and use an inverse Anscombe transform after denoising. They use an adaptive weights smoothing method [42] with adaptive windows to perform the denoising on 3D confocal images. They use the I-divergence (see section 3 for a definition) normalized with the maximal occuring value to estimate the convergence of the denoising. This method reduces the Poisson noise and is a good pre-processing before any deconvolution.

2.2 The point spread function in confocal microscopy

Confocal microscopy is a 3D imaging technique. The physics of the system generates a 3D PSF, thus we model 3D PSF. The PSF can be physically modelled or directly measured. We present here some formulations of the PSF with different approximations.

Diffraction approximation The simplest confocal PSF we can model is the following approximation: we suppose that the confocal PSF is the square of the conventional fluorescence PSF. This is mostly verified for a confocal microscope used in reflexion mode[3] or for a confocal fluorescence microscope with the same excitation and emission wavelengths [29] (see Appendix B). In fluorescence mode, the emission and excitation wavelengths are generally not the same, and the PSF is best modelled as:

$$h(X,Y,Z) = \left|\hat{P}_{\lambda_{em}}(X,Y,Z)\right|^2 . \left|\hat{P}_{\lambda_{ex}}(X,Y,Z)\right|^2 \tag{1}$$

where $\hat{P}_\lambda(x,y,z)$ is the pupil function [50, 51] for a wavelength λ, which is defined in section 3. The 2D pupil function is a diffraction-limited aperture. Due to the wave nature of light, this transforms one single point to a small patch of light, presenting some oscillations: it is known as an Airy diffraction pattern [4]. This generates the XY blur. The Z blur is also due the nature of the light in the third dimension, because of the diffraction-limited optical system. In the Z dimension, authors often relate it to the optical slice thickness or the axial resolution [5, 49, 72].

A more precise model To improve again the PSF model, we can model the pinhole (see Appendix A) as in [62]:

$$h(X,Y,Z) = \left|A_R(X,Y) * \hat{P}_{\lambda_{em}}(X,Y,Z)\right|^2 . \left|\hat{P}_{\lambda_{ex}}(X,Y,Z)\right|^2 \tag{2}$$

[3] "Confocal reflection microscopy (CRM) provides information from unstained tissues, tissues labeled with probes that reflect light, and in combination with fluorescence. [...] A major attraction of CRM for biomedical imaging is the ability to image unlabeled live tissue." in `S. Paddock`, Confocal Reflection Microscopy: The "Other" Confocal Mode, *Bioimaging*, **32**(2):274-278, 2002. CRM uses confocal system with a single laser beam; there is no excitation, nor emission light, only a scanning laser that could be reflected back to the detector. The closer conventional microscope to this scanning mode is a light reflexion microscope, observing opaque object by lighting them from the top.

where A_R characterizes a pinhole with a radius of R, mathematically defined as $A_R(X,Y) = 1$ if $R^2 < X^2 + Y^2$ and 0 elsewhere. The product $*$ denotes the 2D convolution over X and Y. We will use this model in this research report (see section 3). It is now easy to obtain Eq. 1 from Eq. 2 in the limit where the pinhole radius R tends to zero. In this case $A_R(X,Y) \rightarrow \delta(X,Y)$.

In [58, 59, 60], van der Voort et al. are describing a PSF model based on electromagnetic diffraction theory. To calculate the overall PSF of a confocal microscope, they have to compute the time-average of the electrical energy distribution near focus, for both excitation and emission. The final PSF is the product of the time-averaged field excitation lens system, and the emission spectrum of the system. For further details, the interested reader should refer to [58, 59].

Computing a 3D PSF from a 3D pupil function Stokseth's [55] 2D approach is the following: for any defocusing of Z, he can compute the defocus path of the wavefront, and then derive the 2D pupil function for this defocusing. The next step is to calculate the corresponding 2D PSF for this defocusing. Another approach, like Arnison's [3] or Schönle's [48], is to directly compute a 3D pupil function, containing all the defocusing, and then use it to compute the 3D PSF. This is suitable for any optical system, in particular for a confocal microscope. A focused 2D pupil function is a thin disk, whereas a focused 3D pupil function is a part of a thin spherical shell. In both cases, coefficient are constant if we are in-focus; but the defocusing introduces some coefficient variations on the defined surface.

Real PSF measurement Another approach is to measure and use the real PSF for a confocal microscope [12, 13] by using small (a hundred of nanometers) fluorescent beads. The image of such small beads are very sensitive to noise, and the beads themselves are not perfectly spherical, and the diameter is slightly different between two beads. As recommended in the software `Huygens`[4], we have to average some PSF measurements with different acquisitions of microspheres of known-diameter, in the strictly same conditions. In `Huygens`, they recommend to acquire an image field containing at least five beads with the same diameter. In [12, 13], the authors also explain that they need to average the PSF over several different beads to get an accurate measurement. Their averaging consists in four steps:

- extracting some stacks of single beads from raw image acquisition;
- selecting the convenient bead for processing (avoiding bead clusters for instance);
- aligning the selected beads together;
- computing the average PSF.

Doing this, it is possible to reduce the noise of the bead images and to limit the effects of size variations among beads.

[4] `Huygens` is a deconvolution software proposed by Scientific Volume Imaging, `http://www.svi.nl`.

2.3 Deconvolution in confocal microscopy

Many deconvolution methods have been used in fluorescence (conventional or confocal) microscopy. A well-known 2D deconvolution algorithm is the one developed by Agard et al. [1, 2]. It is very fast because it uses only 2D Fourier transforms. Agard initially designed it for widefield microscopy. He makes the assumption that the main part of the blur that appear on an optical section is only due to the nearest section contribution. He uses a theoretical PSF model developed by Stokseth [55]. In [6], Castleman presents Agard's method and apply it with his own widefield PSF model.

In fluorescence confocal microscopy, many 3D deconvolution algorithms have been presented. One of the most common ones is Richardson-Lucy [31, 44] algorithm: this is an expectation maximization algorithm which computes the maximum likelihood estimate (EM-MLE). It is used with different regularization terms; other famous ones are the iterative constrained Tikhonov-Miller (ICTM) algorithm [63, 67] and Conchello's algorithm [9]. Each method assumes that we know the real PSF or an approximation (a model) of the real PSF. For blind deconvolution in 3D fluorescence microscopy, the interested reader should refer to [34, 35]. In these papers, the authors are using an EM algorithm (Richardson-Lucy) with an unknown PSF.

We are presenting here the main deconvolution algorithms that are used in 3D confocal fluorescence microscopy.

Bayesian approach and Richardson-Lucy algorithm Confocal microscopy is a low-photon imagery technique. Light emission is well-approximated by a Poisson process, and so is the 3D image. If an object o is observed as an image i through an optical system with a PSF h and degraded by a Poisson noise, we can write the likelihood as [30]:

$$p(i \mid o) = \prod_x \left(\frac{\left[(h * o)\,(x) \right]^{i(x)} e^{-(h*o)(x)}}{i(x)\,!} \right) \tag{3}$$

Maximizing this probability with respect to o and using a multiplicative gradient-based algorithm leads to RL algorithm (see Appendix C for derivation):

$$o_{k+1}(\mathbf{x}) = \left\{ \left[\frac{i(\mathbf{x})}{(o_k * h)\,(\mathbf{x})} \right] * h(-\mathbf{x}) \right\} . o_k(\mathbf{x}) \tag{4}$$

This algorithm has the property of non-negativity: if the first estimate is non-negative, the further estimates are non-negative. In the presence of noise, the algorithm converges to a solution dominated by the noise, which is not suitable: the noise is amplified after several iterations [62]. But it is nevertheless usefull to use RL because it ameliorates the images both qualitatively and quantitatively before it amplifies the noise [12].

It is also possible to use non-regularized RL with pre-denoised images. In [12], the authors are using a wavelet denoising algorithm before using RL with no regularization; in [28] Kervrann et al. propose an adaptive weight smoothing method for denoising. As the second step, they plan to use a RL algorithm to deconvolve the images.

Richardson-Lucy algorithm with Tikhonov-Miller regularization To force RL to converge to a suitable solution, we can add a prior model on the object, which introduces a regularization term. This a priori functional is designed to avoid noise amplification. Instead of maximizing only the likelihood probability, we maximize the a posteriori probability $p(i \mid o) = p(o \mid i)p(i) \,/\, p(o)$ to take into account the regularization (a priori) term. Tikhonov-Miller (TM) [62, 64, 66] regularization based algorithms are the mainly used algorithms in 3D image restoration. Its expression is given in Appendix C. RL regularized with this functional gives (see Appendix C for derivation):

$$o_{k+1}(\mathbf{x}) = \left\{ \left[\frac{i(\mathbf{x})}{(o_k * h)(\mathbf{x})} \right] * h(-\mathbf{x}) \right\} . \frac{o_k(\mathbf{x})}{1 - 2\lambda_{TM} \triangle o_k(\mathbf{x})} \tag{5}$$

Conchello et al.[5] [9] have derived another form of this algorithm which incorporates TM regularization into RL algorithm for parametric blind deconvolution: the algorithm estimates both the image and the PSF.

Richardson-Lucy algorithm with maximum entropy regularization In Boutet de Monvel et al. [12, 13], the authors use a RL algorithm with maximum entropy regularization (which is clearly explained in [12]). The algorithm scheme is the following:

$$o_{k+1}(\mathbf{x}) = \left\{ \left[\frac{i(\mathbf{x})}{(o_k * h)(\mathbf{x})} \right] * h(-\mathbf{x}) \right\} . o_k(\mathbf{x}) - \lambda_{entropy} . o_k(\mathbf{x}) . \ln\left(o_k(\mathbf{x})\right) \tag{6}$$

The regularization parameter is $\lambda_{entropy}$. As reported in [12, 13], for high values of $\lambda_{entropy}$, the algorithm of Eq. 6 becomes unstable.

The iterative constrained Tikhonov-Miller algorithm Another famous algorithm is the iterative constrained Tikhonov-Miller algorithm (ICTM). This algorithm finds iteratively the minimum of the TM constraint [57] in the case of Gaussian noise. The non-negativity constraint is ensured by setting to zero each negative values at each iteration [62, P.40]. In [60, 62], they consider the image formation model: $i = o * h + n$ where i is the image, o the object, h the PSF and n a Gaussian noise. The non-iterative form of TM regularized inversion has to fulfill several criteria. Defining the norm $\|.\|^2$ as:

$$\|f\|^2 = \int\int\int |f(x,y,z)|^2 \, dx \, dy \, dz \tag{7}$$

the first criterion is $\|i - h * o\| \leq \varepsilon$ with ε an estimate of the noise. The second criterion is used to exclude wildely oscillating functions, and to stabilize the functional: $\|o\| \leq \mathcal{E}$, $\mathcal{E}$ being an energy bound. The non-iterative algorithm have to minimize the following functional:

$$\Phi(o) = \|i - h * o\|^2 + \left(\frac{\varepsilon}{\mathcal{E}}\right)^2 \|o\|^2 \tag{8}$$

[5] Conchello, Preza [43] and McNally have developed a deconvolution freeware known as `Xcosm` which is available for 8-bits X servers on several Unix platforms. See `http://www.essrl.wustl.edu/~preza/xcosm/`.

Using capital letters to denote Fourier transform, and minimizing the functional $\Phi(O)$, the unconstrained TM solution is:

$$O = \frac{H^*.I}{\|H\|^2 + \left(\frac{\varepsilon}{\mathcal{E}}\right)^2} \tag{9}$$

This algorithm requests a modest computational effort, but has to be used with images acquired under very precise conditions to estimate the parameters (ε and $\mathcal{E}$) accurately.

ICTM algorithm, used with non-negativity constraint, is less sensitive to errors than non-iterative TM restoration. But it is also more computationally consuming. It is based on iterative conjugate gradient optimization. The ICTM algorithm is computed as:

$$\hat{o}_{k+1} = \hat{o}_k + \beta.p_k \tag{10}$$

β is a stepping distance, and k is the iteration number. For each iteration, the non-negativity constraint is applied. The functional p_k is recursively defined by:

$$p_k = r_k + \frac{\|r_k\|^2}{\|r_{k-1}\|^2}.p_{k-1} \tag{11}$$

The steepest descent r_k is the negative gradient of the functional Φ:

$$r_k = -\frac{1}{2}\nabla_o \Phi\left(\hat{o}\right) = \lambda_{ICTM}\hat{o}_k + h * h^* * \hat{o}_k - h^* * i \tag{12}$$

λ_{ICTM} being a parameter containing the term $\left(\frac{\varepsilon}{\mathcal{E}}\right)^2$.

The Carrington algorithm In [62, 66, 67], van Kempen et al. are using a derivate of the ICTM, known as the Carrington algorithm. This algorithm also minimizes the Tikhonov functional under the constraint of non-negativity and a finite support [62]. The method is not too much noise-sensitive. For a derivation, see [62, p.42].

2.4 Total Variation

Total Variation (TV) is a well-known method for denoising images. It has been introduced by Rudin, Osher and Fatemi [45]. This functional has a $\mathcal{L}^1$ norm and its effects are to smooth homogeneous regions while preserving edges. On the contrary, a $\mathcal{L}^2$ as TM tends to smooth edges. Nevertheless, even if TV preserves the contours, one drawback of the method is to round the corners [16].

2D Total Variation in image processing The Total Variation norm (TV) was first described in [45] as an iterative denoising method. It is non-quadratic and preserves edges in blurred and noisy images. According to [45], the 2D TV semi-norm is given by:

$$\mathrm{TV}(o) = \int |\nabla o| \, dx \, dy = \int \sqrt{\frac{\partial o}{\partial x}^2 + \frac{\partial o}{\partial y}^2} \, dx \, dy \tag{13}$$

The discrete form of this equation is $\Sigma_{l,m}\sqrt{(o_{l+1,m} - o_{k,m})^2 + (o_{l,m+1} - o_{k,m})^2 + \eta^2}$. It needs an arbitrary parameter $\eta < 10^{-3}$ to ensure differentiability at points where $(o_{l+1,m} - o_{k,m})^2 + (o_{l,m+1} - o_{k,m})^2 = 0$ (see Appendix D and [41] for the discrete form of the 3D TV semi-norm). In [16, 45], the authors use the hypotheses of a Gaussian noise to minimize the following functional:

$$E_{TV}(o) = \int |\nabla o| + \lambda \int |h * o - i|^2 \tag{14}$$

where h is the PSF, o the object and i the measured image, λ being the regularization parameter. With the Rudin et al. denoising approach, smaller details, such as textures, are destroyed if λ is too small [37]. In [21], the authors propose to use an adaptative fidelity term to preserve textures. They impose a spatially varying fidelity term which modulates the amount of denoising over image regions.

To our knowledge, TV is used in 2D for noisy images restoration, mostly in presence of an additive noise [45]. TV is very efficient for non-textured objects, because it smoothes homogeneous regions and keeps sharp edges, but the performance on textured images are poor. To avoid texture loss, Malgouyres [32, 33] used an approach combining TV for smoothing, and wavelet decomposition to preserve small details. Moreover, he shows that the combination of these two methods avoids ringing (from wavelets) and staircasing (from TV). The degraded image model is a 2D image convolved with a low-pass filter, and corrupted by Gaussian noise. Another proposed approach is to solve this problem by computing an adaptive regularization parameter [21].

3D Total Variation in tomographic image restoration In the domain of 3D tomography, Charbonnier [7] used a similar method to reconstruct 3D images in the presence of Poisson noise. He used a Richardson-Lucy algorithm regularized with Total Variation but incorporated the TV in a different way. Another difference with our method is also that the PSF is not normalized as ours.

On simulated images which present several homogeneous regions, the algorithm gives good restoration results after a hundred iterations. Charbonnier applied this algorithm on synthetic data containing ellipses, and he observes a good restoration but a loss of small object after the restoration, due to the regularization.

Still in 3D tomography, Persson et al. [41] present an iterative Bayesian reconstruction algorithm based on three-dimensional TV semi-norm. This is a 3D extension of Rudin et al.

[45] 2D iterative noise removal algorithm. Persson et al. assume Poisson statistics for their data. Their algorithm was designed in the limited view angle case.

Concerning the 3D implementation of the TV, more details can be found in Appendix D of this document, where we clearly write the derivative form of the TV we are using in the deconvolution method presented in this research report.

2.5 Summary

We have presented here the works related to ours. In confocal microscopy, denoising could be sufficient to obtain degradation-free images. But if the blur is important, one have to take into account the PSF of the system: the PSF can be approximated knowing the physics of the optical system, or measured using methods that are very precise. An other method is to estimate this PSF during the image restoration process. By knowing this PSF or by estimating it, we are able to perform 3D image deconvolution.

For our purpose, the most interesting algorithms we have presented are those using iterative RL and a regularization constraint. We have presented a TM constraint [62, 64, 66] and a maximum entropy constraint [12, 13]. We remind that we are presenting in this research report an algorithm based on RL with a new regularization term based on Total Variation.

3 The proposed deconvolution method

We model the blur as the convolution of the object with the point spread function (PSF) of the microscope. The noise is mainly shot noise due to low-photon imagery, modelled as Poisson statistics. We present here the deconvolution method we are using: the well-known Richardson-Lucy multiplicative algorithm. Since it does not converge to a noise-free solution (it amplifies the noise after several iterations), we regularize it using a functional derived from the Total Variation.

We then present some quantitative criteria, as I-divergence and mean square error, to quantify the results. The criteria are not used the same way if we are dealing with real data or with simulated objects. The last point of this section is about the criterion that stops the iterative algorithm.

3.1 The image formation model

In this section, we present the image formation process for a confocal microscope. We first describe the global model for image formation: the point spread function (PSF) which characterizes the optical system, and the noise model (Poisson noise).

The image statistics Confocal microscopy is a low-photon imagery technique. To obtain a sufficient number of photons, a photomultiplier tube (PMT) converts the light to an electrical impulse, that is stored as a discrete value after a A/D conversion [28, 39]. Because each detected photon gives an electrical impulse, the PMT works as a photon counter, and the image statistics is best described by a Poisson process. The noise model is a Poisson multiplicative noise, or shot-noise.

When acquiring real images with a confocal microscope, an averaging can be performed during acquisition: the acquisition process is done point by point, and then line by line. Averaging is done by acquiring two or more times the same line, and by averaging all the values. Doing this changes the image statistics, but the Poisson model is still valid. The higher the averaging, the more the statistics tends to Gaussian statistics.

The global image formation model for a confocal microscope For any optical system, the image formation model is described by combining a PSF convolution and a noise degradation on a perfect undegraded image. According to [47, 62, 65], the most used image formation model in confocal microscopy is:

$$i = \wp(o * h + b) \tag{15}$$

$\wp$ is a Poisson noise, i the observed image, o the object, h the PSF and b the background signal. We choose a well-known Confocal Laser Scanning Microscope (CLSM) image formation model, under the assumptions of a translation invariant PSF, incoherent imaging and monochromatic light (excitation and emission). CLSM is a low-photon imagery technique, therefore the noise obeys a Poisson multiplicative law.

As a first approximation in the model we use, we omit the modelling of the background. With no background estimation, the image formation model with the same notations as Eq. 15 becomes:

$$i = \wp(o * h) \tag{16}$$

In section 4.4, we discuss about the use of a background term in the model, and for our purpose we show on results that there is no improvement by modeling a background.

3.2 The confocal point spread function

We discussed about PSF measurement and modelling in section 2.2. For our purpose, we are using a PSF model that is used both for the creation of simulated images and for the deconvolution. It has been already used in [62], and is derived from [53]. For a defocusing Z, the 3D PSF [53] is given by:

$$h(X,Y,Z) = \left| A_R(X,Y) * \hat{P}_{\lambda_{em}}(X,Y,Z) \right|^2 . \left| \hat{P}_{\lambda_{ex}}(X,Y,Z) \right|^2 \tag{17}$$

with the same notation as in section 2.2. The excitation (resp. emission) wavelength is λ_{ex} (resp. λ_{em}). A_R denotes a circular aperture (as in Eq. 2), and $\hat{P}_\lambda$ is the 2D Fourier transform on X, Y of the pupil function P_λ, given by [4, 19, 22]:

$$P_\lambda(U,V,\,Z) = \Pi_\rho \left(\sqrt{U^2+V^2} \right) . e^{\frac{2i\pi}{\lambda} . W_{lpe}(U,V,Z)} \tag{18}$$

Here the complex term $W_{lpe}(U,V,\,Z) = \frac{1}{2} Z \,(1 - \cos 2\alpha)$ is the aberration phase[6] derived from [22]. Its general form is complex, the real part describing the default of focus and its imaginary part the spherical aberration of the lens, after Stokseth [55]. Here W_{lpe} is real, as we do not take spherical aberrations into account. $\rho = \frac{NA}{2\lambda}$ is the lateral cut-off frequency, and $NA = n_o \sin\alpha$ the numerical aperture which is related to the amount of light entering the microscope (see Fig. 1) and the immersion medium refractive index n_o. The phase W_{lpe} depends only on the defect of focus. This theoretical model of the confocal PSF does not take into account geometrical (e.g. spherical) lens aberrations and refractive index induced diffraction.

The model takes into account the finite size of the pinhole used for the image acquisition. The pinhole is represented here by the distribution $A_R(X,Y)$. We have represented the PSF model in Fig. 2.

3.3 The multiplicative Richardson-Lucy algorithm

We summarize here how to obtain the RL algorithm (for more details see Appendix C). A common approach to image restoration uses a probabilistic framework. One way to estimate

[6] "lpe" stands for "length path error". This is the wavefront difference of path between a focused beam and a defocused one [4, 55].

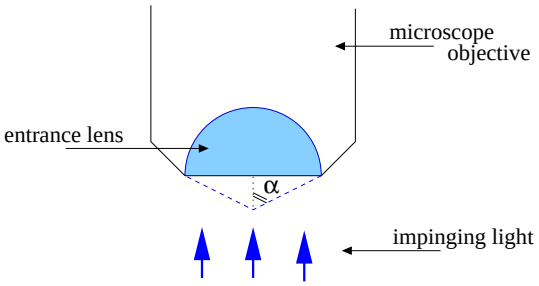
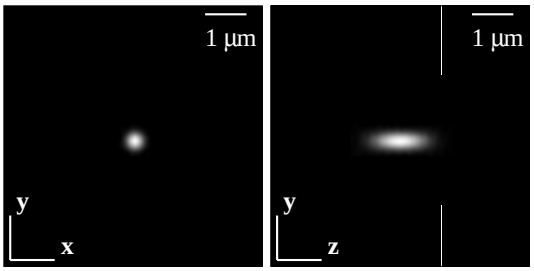


Figure 1: The numerical aperture $NA = n_o \sin\alpha$ the numerical aperture which is related to the amount of light entering the microscope. It is the half angle of the maximum ray of light that could enter the objective according to Snell's law [4, p.40].

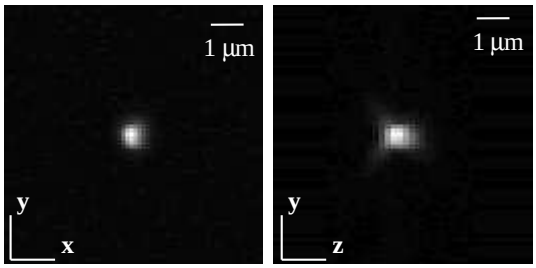


(a) (b)

Figure 2: The PSF calculated with the model defined on Eq. 17. We represent it with the same sampling in each direction to reveal the ratio of 1 : 3 usually observed for a confocal microscope. We use this PSF for all the computations we are presenting here, for blurring synthetic objects, and for deconvolving both synthetic and real data. For the computations we use it with the sampling values of the considered image stack. To simulate this PSF, we used the values defined in section 5.1.

the initial image (before degradations occur) is to maximize the likelihood probability $p(i \mid o)$. In CLSM, the detected light can be modelled as a Poisson process. We suppose that the noise is independent from one pixel to another. The recorded image is also a Poisson process, and the likelihood is then:

$$p(i \mid o) = \prod_{\mathbf{x}} \left(\frac{[(h * o)(\mathbf{x})]^{i(\mathbf{x})}\, e^{-(h*o)(\mathbf{x})}}{i(\mathbf{x})!} \right) \tag{19}$$



(a) (b)

Figure 3: An approximation of the real PSF. Two views of the image of a small bead of diameter 175 nm: (a) is the lateral XY view, (b) the axial XZ view. Initially, the size of a voxel is 89x89x230nm in X, Y and Z. In the figure, it is now 89nm in each of the three directions. Note that there was no resampling. This is the best approximation we have to illustrate the real PSF; the lateral size is 491 nm (6 pixels). and the axial size is 1150 nm. An good estimation of the size of the PSF is to remove this approximate size of 175 nm of the bead, that gives a PSF with a size of 316 nm in XY and 975 nm in XZ.

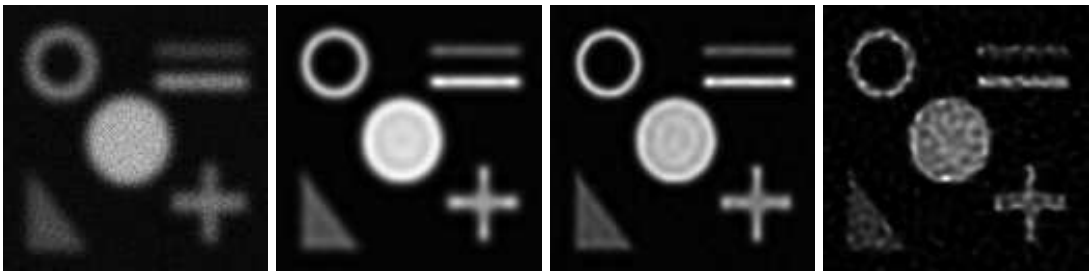
The non-regularized Richardson-Lucy [31, 44] algorithm minimizes the functional $-\log p(i \mid o) = J_1(o)$ (up to an additive constant), giving the maximum likelihood (ML) estimation:

$$J_1(o) = \sum_{\mathbf{x}} \left(-i(\mathbf{x}) \log \left[(h * o)(\mathbf{x})\right] + (h * o)(\mathbf{x})\right) \tag{20}$$

RL is a multiplicative gradient-based iterative algorithm. Let $o_k(\mathbf{x})$ be the estimate at iteration k. Then one RL iteration is given by:

$$o_{k+1}(\mathbf{x}) = \left\{ \left[\frac{i(\mathbf{x})}{(o_k * h)(\mathbf{x})} \right] * h(-\mathbf{x}) \right\} . o_k(\mathbf{x}) \tag{21}$$

Let us note that Eq. 21 can also be obtained by deriving an expectation maximization (EM) algorithm for the ML estimation (Eq. 19). An important property of this scheme is that it ensures non-negativity if the first estimate (here taken as a constant 3D image, with a value equal to the mean of the image stack) is non-negative. However, because the inversion problem is ill-posed and the ML estimator is non-regularized, the solution given by Eq. 21 when $k \rightarrow +\infty$ is only noise, due to noise amplification during the inversion process. To obtain good results, the algorithm is stopped before the noise amplification. In order to improve the trade-off between a good deconvolution with sharp edges and noise amplification, we propose to regularize the solution by minimizing its TV. Fig. 4 shows the evolution of the standard RL deconvolution: the estimate seems to converge to a non-optimal solution according to us, and according to the I-divergence criterion.



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

Figure 4: Standard RL does not converge to a suitable solution. Each image (a)-(d) shows one section of a synthetic 3D image stack. (a) degraded image before any deconvolution (I-divergence is 6.11); (b) the best deconvolved image after 17 iterations (I-divergence is 1.36); (c) after 100 iterations, some intensity rebounds inside the objects are noticeable (I-divergence is 1.45); (d) after 4075 iterations, many intensity oscillations are visible (I-divergence is 6.19). This figure shows the need of a regularization term to avoid.

3.4 Total Variation regularization

Here we choose a regularization by Total Variation, as in [45]. It results in minimizing the sum of J_1 and the Total Variation J_{reg} (See Appendix C for the derivation.):

$$J_1(o) + J_{reg}(o) = J_1(o) + \lambda \sum_{\mathbf{x}} |\nabla o(\mathbf{x})| \tag{22}$$

Using the $\mathcal{L}^1$ norm over ∇o rather than the $\mathcal{L}^2$ norm as in Tikhonov regularization [8, 66] allows to impose edge preserving smoothing of the solution. It can be shown that the smoothing process introduced by J_{reg} acts only in the direction tangential to level lines (edges) and not in the orthogonal direction, so edges are preserved. The derivative of J_{reg} w.r.t. o is a nonlinear term $\frac{\partial}{\partial o} J_{reg} = -\lambda div\left(\frac{\nabla o}{|\nabla o|}\right)$ where div stands for the divergence. We minimize Eq. 22 using the multiplicative gradient-based algorithm (or equivalently by using EM algorithm for the penalized criterion of Eq. 22), and we adopt an explicit scheme, as in [23], defined by:

$$o_{k+1}(\mathbf{x}) = \left\{ \left[\frac{i(\mathbf{x})}{(o_k * h)(\mathbf{x})} \right] * h(-\mathbf{x}) \right\} . \frac{o_k(\mathbf{x})}{1 - \lambda\, div\left(\frac{\nabla o_k(\mathbf{x})}{|\nabla o_k(\mathbf{x})|}\right)} \tag{23}$$

3.5 Quantification of the results

To quantify the quality of the deconvolution we use the I-divergence [11] and the mean square error (MSE) criteria. According to [11], the I-divergence is the only consistent measure in

the presence of a non-negativity constraint. The I-divergence between two 3D images A and B is given by:

$$I_{A,B} = \sum_{ijk} \left\{ A_{ijk}.\ln\left[\frac{A_{ijk}}{B_{ijk}}\right] - (A_{ijk} - B_{ijk}) \right\} \tag{24}$$

while the mean square error is given by:

$$MSE_{A,B} = \sum_{ijk} \{A_{ijk} - B_{ijk}\}^2 \tag{25}$$

We can notice that the I-divergence is non-symmetric: $I_{A,B} \neq I_{B,A}$, but MSE is.

Quantifying simulated data For simulated data, we simulate an object, and then we add some degradations (blur and noise) according to physical assumptions described in section 3.1. We have two advantages by working with synthetic degraded images:

- we know exactly the initial object we want to estimate. Then, it is easy to quantify results by comparing them to the initial undegraded object.
- the degradation model is exactly known: we perfectly know the PSF and the noise statistics obeys a Poisson law.

Thus, to quantify the results, we compute the I-divergence and the MSE values by comparing the final estimate to the initial undegraded object. In this case, both criteria have to decrease with iterations. Ideally, a perfect deconvolution should end with I-divergence and MSE both equal to zero. A typical final value for I-divergence is around 10^{-5} if it is normalized to the object. Here the I-divergence is not normalized to the object.

Quantifying real data We have problem for real data because we do not know exactly what the initial object before degradations was, as we only know it after the image acquisition (and the degradations). In this case, we can only compare the estimate to raw data. But it is not as well adapted as in for simulated data:

- the image restoration generates an image that is more or less free of noise and blur. The I-divergence and the MSE values between this estimate and the raw data are positive, and strictly increasing.
- a high value of these criteria could correspond to a good restoration or a very poor one: the only thing which is measured is the distance between the estimate and the initial object.

`Huygens` computes after each iteration a "quality factor", which is based on I-divergence [61].

Ending the iterations To stop the iterations, we define a difference measure between two successive iterations. If the difference is smaller than a threshold, we stop the computation, and we suppose that the last estimation is the best one. The criterion is defined as the following:

$$\chi_{k+1} = \frac{\Sigma_{x,y,z=1}^{N_{x,y,z}} |o_{k+1}(x,y,z) - o_k(x,y,z)|}{\Sigma_{x,y,z=1}^{N_{x,y,z}} o_k(x,y,z)} < t \tag{26}$$

Typical threshold values t used are around 10^{-4} or 10^{-6}, depending on the precision we want. It also depends on the regularization parameter. For high values of the regularization parameter, the stop criterion should not be too small, because we noticed that the χ_k could oscillate between two values.

3.6 Summary

We have presented the point spread function model we are using, which is based on physical theory of the image formation in a confocal microscope. We use this model to simulate the image formation on simulated data, and then evaluate the results after deconvolution. On real data, we use the PSF as an approximation of the real one to deconvolve real images, using all the physical parameters that are involved in the PSF model (light wavelengths, pinhole diameter, objective numerical aperture). We have presented an algorithm based on Richardson-Lucy with Total Variation regularization, and also the quantitative criteria that are used to quantify the results, for real and simulated data.

4 Results on simulated data

In this section, we present some results on simulated data. In section 5, we will present results on real data. It is more convenient to begin with simulated data, because we know exactly the point spread function (PSF) for degradation simulation and deconvolution, and we are able to use quantitative criteria to measure a distance between the original data and the estimate.

4.1 Operating mode

As in [14], we use the following scheme to perform the validation of the method on simulated data:

- Step one: create the object. We generate a 3D object that has some specific characteristics: smoothness, corners, borders, textures, fine structures,... and then, we add a constant background. The simulated images are 3D images created with some default brushes of `The GIMP`[7]. We have both geometrical (stars, crosses) and non-geometrical objects. Some are smooth, others have fine structures.
- Step two: simulate object degradations. We simulate the point spread function (PSF) at a chosen resolution, and we blur the initial object. After that, we add Poisson noise. We obtain a degraded object which can be used to test the deconvolution algorithm.
- Step three: perform the deconvolution. We know exactly the PSF and use it to perform the iterative deconvolution. The algorithm ends when the difference (see paragraph 3.5) between two consecutive image estimates falls below a defined threshold.
- Step four: qualitatively and quantitatively interpret the result. Since we know exactly the original object, it is easy to compare qualitatively the final estimate to the original non-degraded image. It is also very easy to define quantitative criteria between original and final estimated images.

We use this procedure for each 3D synthetic image presented in this section.

4.2 Results on 3D simulated data

Here we present experimental results of the proposed deconvolution model on 3D simulated data. These results have been originally presented in [14]. Fig. 5 represents three different objects. These objects (a)-(c) are artificially blurred using the image formation model (see section 3), and corrupted by a Poisson noise to give the degraded images in fig. 5 (d)-(f). Deconvolution results using standard RL are shown in fig. 6 (a)-(c). For standard RL, we present the results for the number of iterations that achieve minimum I-divergence. The algorithm amplifies the noise after that point.

[7] See `http://www.gimp.org`

object type	cylinder	composed	sphere
I-divergence for standard RL	0.766	1.365	1.599
I-divergence for RL and TV	0.220	0.691	0.725
improvement (I-divergence)	71.3 %	49.4 %	54.7 %
MSE for standard RL	169.965	262.813	411.727
MSE for RL and TV	49.225	100.412	212.977
improvement (MSE)	71.0 %	61.8 %	48.3 %

Table 1: I-divergence (top) and MSE values (bottom) for various objects (see Fig. 5). The first row gives object type; the second row is the minimal value obtained with non-regularized RL; the third row is the value obtained at convergence of the regularized method with TV; the fourth row gives the improvement obtained by RL with TV regularization compared to standard RL. Fith and sixth rows gives the MSE value for standard RL and RL with TV. Last row shows the improvement between standard RL and RL and TV.

The deconvolution results using RL with TV regularization are represented in the second row (d)-(f) of Fig. 6. Even if RL regularized with TV is theoreticaly always convergent, numerically we noticed that the regularization parameter λ_{TV} should not be too small (less than 10^{-6}, RL is dominated by the data model), nor too large (of order unity, RL is dominated by the regularization term). If λ_{TV} is too large, the denominator of Eq. 23 can become zero or negative: this must be avoided because small denominators create points of very high intensity, that are amplified at each iteration. A negative value violates the non-negativity constraint. We use a parameter $\lambda_{TV} = 0.002$ for computations and iterations stop when the difference between two images is less than a chosen threshold (10^{-5}).

For each simulation, we choosed the PSF represented in fig. 2 to blur original objects and then to deconvolve the degraded data. We can compare the two methods by considering the deconvolved images in fig. 6 (a)-(c) and (d)-(f): the images (d)-(f) obtained with RL regularized with TV are smoother (oscillation artefacts almost absent) than (a)-(c), obtained with standard RL; there is no noticeable noise, only some weak variations in the background; the images are sharper at the borders, and finally, the sizes of the objects are closer to those of the non-degraded images. Compare, for instance, axial and lateral views in subimages (b) and (e). Table 1 presents the improvement in restoration quality. As measured by I-divergence (resp. MSE), this improvement ranges from 49 to 71% (resp. 48 to 71%) depending on the image.

4.3 Results on 3D fine structures and textured surfaces

Fine structures often exist in real biological specimens. We study the limits of the proposed method on such simulated fine structures before applying the method on real data. We represent these structures as lines of width of order of a pixel.

In Fig. 7, we have represented a 3D synthetic object combining texture and fine structure (a). The degradation (b) is sufficiently strong to almost hide these details. Standard RL (c)

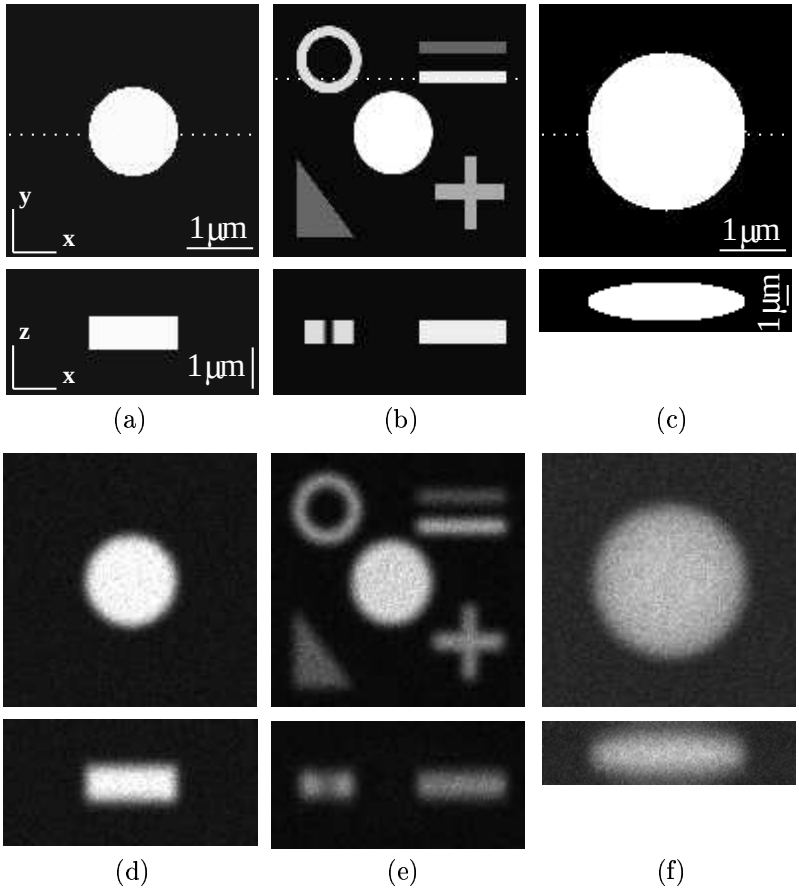


Figure 5: Simulated test objects and degraded images. First row represents original synthetic objects, and second row blurred and noisy objects. First and second columns represent 128x128x64 test objects with voxel size 250x250x600nm in X, Y and Z. Dotted lines in XY images (top row) show where the YZ cut (bottom row) is taken. (a): cylinder with intensity of 250 and background of 20. (b): composed object with different intensities: 255 for the cylinder, 221 for the annulus, 170 for the cross, 102 for the triangle, 238 and 102 for the equal sign, background 10. (c): sphere with intensity of 200 and a background of 40. Same lateral scale but axial scale is double. (d)-(f): the degraded objects.

gives very good results if we stop it before it amplifyes the noise: we recover the fine structure and a part of the texture. In Z, there is still some blur. With the deconvolution using RL with TV, the blur is almost completely removed, and the fine structure is discernable again.

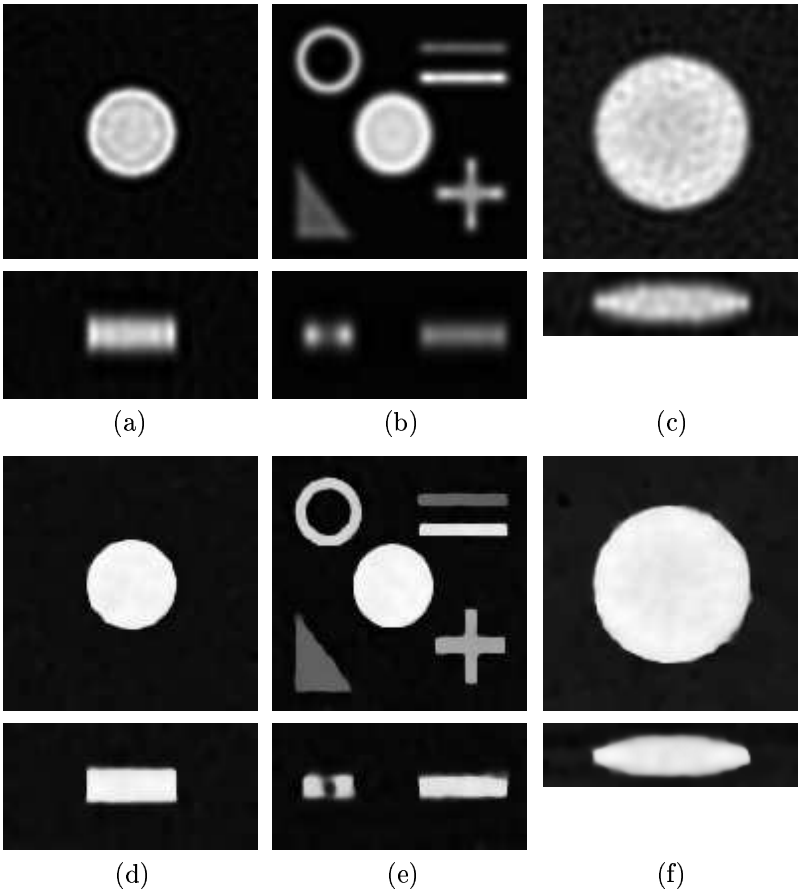
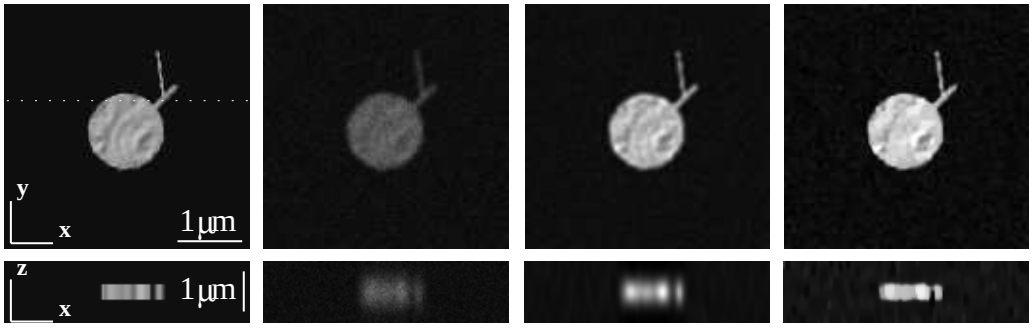


Figure 6: Deconvolution with and without TV regularization. Deconvolution of synthetic test images using standard RL (first row) and RL with TV regularization (second row) with $\lambda_{TV} = 0.002$. First row: images deconvolved with standard RL. There is still some intensity oscillations and some blur; second row: images deconvolved with RL and TV regularization. The homogeneous regions are smoothed and the borders are sharp.

Concerning the texture, however, we see one drawback of TV regularization on the XY section: a stair-casing effect. The quantitative values of I-divergence are quite similar, and even if RL with TV visually gives a better result by improving the contours, standard RL results are acceptable.



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

Figure 7: Deconvolution of a textured image with fine structure. Deconvolution of synthetic test images using RL with $\lambda_{TV} = 0.002$. First line represents the lateral XY view of the image (center of the stack), and the second line represents a YZ view (taken were dotted lines are figured in (a)). (a): original synthetic image; (b): blurred and noisy images degraded as described in section 3; (c): images deconvolved with standard RL (I-divergence is 0.462); (d): images deconvolved with RL regularized with TV (I-divergence is 0.404).

4.4 Background estimation

In [65], the authors are studying the influence of the background estimation on the quality of the results and the speed of the convergence. In our experiments, the speed of the convergence is improved with $b = 0$, but the restoration results are the same with or without background estimation.

We have used a simple method of background estimation. The background estimation we use is only computing one histogram, and finds the first maximum of this histogram; the first maximum corresponds to the mean value of the background. We substract this estimated value 3D image and set negative value to zero. It is valid with (real or simulated) images that fullfill the following hypothesis: the background volume must be the most important class of the 3D image in term of voxel number. It is often the case for small specimen observations, and it will be always the case in the present report. Under these hypothesis, this is equivalent to find the grey-level global maximum of the satck.

This imposes a strong support constraint[8], that we are not using here. We present in fig. 8 the results of the deconvolution on a 3D simulated image stack. On the original object, the modelled background is 8, and on the degraded data, we compute a mean background of 10, which is the maximum value to remove. We can see in fig. 8 the original object (a), the degraded object (b), and the results of the deconvolution (Richardson-Lucy with Total Variation) (c) without and (d) with background estimation and substraction. Our purpose

[8]A support constraint is a constraint on the spatial localization of the object. It generally supposes that everything outside this support is zero.

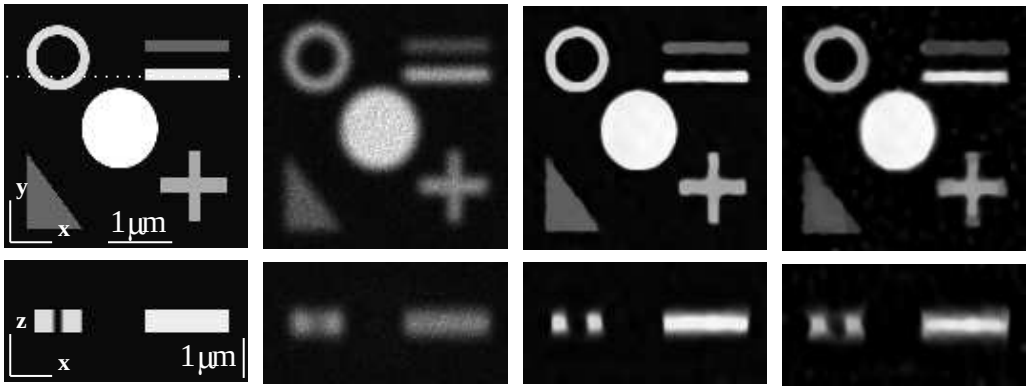


Figure 8: The influence of the background estimation on the deconvolution using Richardson-Lucy with Total Variation regularization. The regularization parameter is fixed to 0.002. First and second columns represent synthetic object (a) and degraded image (b) simulated using Eq. 16 ; third and fourth columns show the results of the deconvolution without background estimation (c) and with background estimation and substraction (d). The computed estimation is 10 (real value is 8). The object is a 128x128x64 object with a voxel size of 250nm in XY and 600nm for YZ. Dotted lines in XY images (top row) show where the YZ cut (bottom row) is taken.

is to compare the quality of the results (c) and (d), with and without background estimation and substraction. Qualitatively, the images are the same in each directions, except for a small difference of contrast: the contrast in part (d) is higher than the contrast in part (c), but we see some oscillations in the background in part (d). Moreover, we can notice a decrease in the intensity of the objects at their extremities. This leads to a high value of the final I-divergence: the I-divergence final value is greater with background removal (around 42) than with no background removal (around 0.69). Quantitatively, the final background is estimated to 3 in both case.

In some other experiments, we tried to remove greater values than those we estimate, and we performed the deconvolution. Even if we remove a background greater by only 1 grey-level, the results are not satisfying: we observe some intensity oscillations at the borders of the objects, that are not suitable.

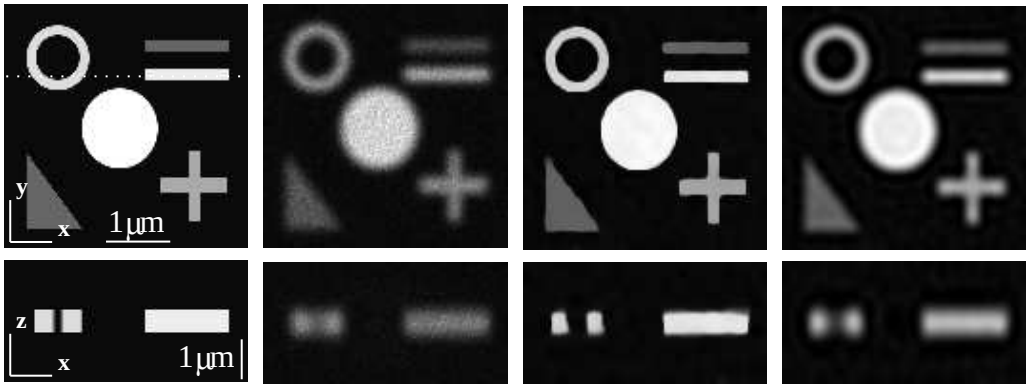
4.5 Tikhonov regularization

If we choose another prior model for the object, the algorithm will be different. One of the most famous regularization term in inverse problems is Tikhonov-Miller (TM) [56, 57] regularization, often used in confocal microscopy [58, 62]. The main difference with TV

regularization (Eq. 23) is the square in the term corresponding to the object model (see Appendix C). TM-based algorithms have some known limitations, the main one being that they do not preserve object edges. Minimizing this functional in the way of obtaining a RL-based algorithm yields[9]:

$$o_{k+1}(\mathbf{x}) = \left\{ \left[\frac{i(\mathbf{x})}{(o_k * h)(\mathbf{x})} \right] * h(-\mathbf{x}) \right\} . \frac{o_k(\mathbf{x})}{1 - \lambda_{TM} \triangle o_k(\mathbf{x})} \tag{27}$$

Here again, we see that if the regularization parameter λ_{TM} is too high, we risk to violate the non-negativity property of the RL algorithm. From experiments we learned that we have to use smaller values for TM regularization parameter than for TV. The chosen values are around 10^{-4}. In Fig. 9, we choose $\lambda_{TM} = 2.10^{-4}$; we see the original objects (a), the degraded objects (b), and the deconvolution results for TV (c) and TM (d).



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

Figure 9: Comparison between Tikhonov-Miller and Total Variation regularization. The regularization parameter was fixed to $\lambda_{TM} = 3.10^{-4}$ for TM and to $\lambda_{TV} = 0.002$ for TV. For TM, this is the highest value we can choose without obtaining negative values in the algorithm. First column represents the synthetic object (a) and second column the degraded object (b); third and fourth columns show the results of the deconvolution with TV (c) and with TM regularization (d). The I-divergence for (c) is 0.69, and for (d) 1.25.

We have to compare third and fourth columns of Fig. 9. The results obtained with TV are better than those obtained with TM: Fig. 9 (d) is more blurred after processing than (c). This is a known effect of the TM regularization that does not respect the object edges and smoothes them. The denoising is well performed, and the sizes of the deconvolved objects are smaller than in reality. TV regularization preserves sharp edges and better respects the sizes of the object, especially on the YZ view. At convergence, the I-divergence value for TM is 1.25 and 0.69 for TV. The stop threshold is fixed to 10^{-5} for Fig. 9.

[9]See Appendix C for derivation.

4.6 Additive algorithm with Gaussian statistics

In this section, we no longer assume that the image statistics is given by a Poisson distribution. Here we suppose that the image statistics is Gaussian and obeys the following image formation model:

$$i(x) = o(x) * h(x) + n_G(x) \tag{28}$$

where n_G is an additive white Gaussian noise. By choosing the Total Variation (TV) as a prior model on the object, we obtain the following algorithm (see Appendix C for derivation):

$$o_{k+1}(\mathbf{x}) = o_k(\mathbf{x}) + \alpha \left[\left(h^* * i\right)(\mathbf{x}) - \left(h^* * h\right) * o_k(\mathbf{x}) - \frac{\lambda}{\alpha} div \frac{\nabla o_k}{|\nabla o_k|} \right] \tag{29}$$

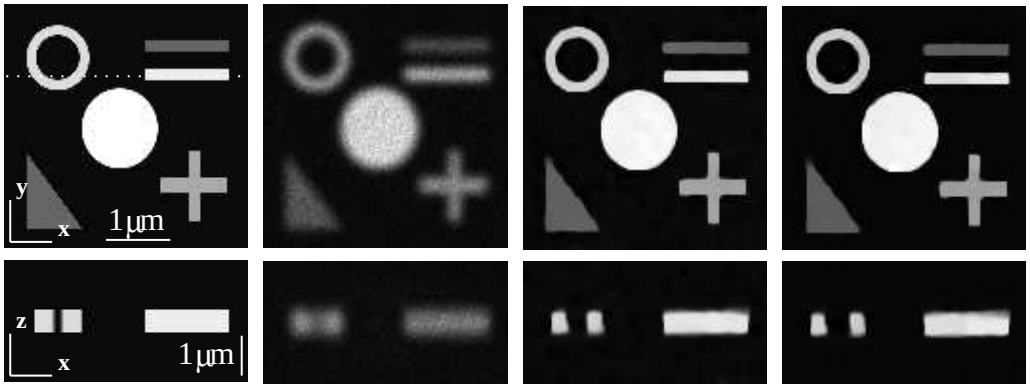
This is an additive gradient-based algorithm with TV regularization. We choose the same regularization parameter of 0.002. In a gradient-descent algorithm, one must choose a sufficiently small step α for the descent to converge slowly to the solution, to limit the risks of stopping at a non suitable (insufficiently restored) solution. We choose a gradient step of $\alpha = 1.0$. The results are shown in fig. 10. One advantage of using a gradient-descent algorithm is that there is no chance to violate the non-negativity constraint (as for RL-based algorithms) because it is not an intrinsic propertiy of the algorithm: we just have to impose ourself a non-negativity by setting to zero at each iteration the negative values of the intensity. We are now able to use large value for the regularization parameter, larger than for RL (see in fig. 10, 0.002 for RL and 0.1 for the gradient-descent algorithm).

The results obtained under the hypothesis of a Gaussian noise are very satisfying: the blur is removed, and there is no intensity oscillations inside the objects. Moreover, the object dimensions are of the same order than in reality (see YZ view of Fig. 10). Even if the simulated blur follows a Poisson law, the deconvolution of the images under the assumption of a Gaussian noise and a TV regularization gives very good results.

4.7 Conclusion

The deconvolution method based on a Richardson-Lucy algorithm with Total Variation regularization gives very good results on 3D synthetic images. The images are degraded using a PSF model and a multiplicative Poisson noise. It is possible to measure the amount of restoration by using the I-divergence between the non-degraded image and the result of the deconvolution. On geometrical object, we can see that the method does not smooth the edges, but rounds the corners (see Fig. 6).

For fine structures and textured objects, we have tested the same algorithm on a synthetic object in fig. 7). We can still notice some noise in deconvolved data (d), but the object itself is well localized. Concerning the texture, we recover some of the larger scale features (compare for instance subimages (a) and (d)), but very fine details of the texture are lost. Moreover, we can notice the well-known stair-casing effects of the TV inside the object. The improvement of the image using RL with TV is very high, but we have to be careful with small details (like texture).



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

Figure 10: Under the assumption of Gaussian noise or Poisson noise. We compare the use of Total Variation regularization for images degraded by a Poisson noise, supposing that the noise is Poissonian (c) or Gaussian (d). First column represents the synthetic object (a) and second column the degraded object (b); third and fourth columns show the results of the deconvolution with TV (c) supposing Poisson noise, and with TV (d) supposing a Gaussian noise. In the first case (c), the deconvolution is processed with a multiplicative algorithm (RL) with TV regularization. In the second case (d), and the deconvolution is performed using an additive gradient-descent algorithm. The regularization parameter was fixed to to 0.002 for RL and to 0.1 for gradient-descent algorithm. Qualitatively, the Gaussian approximation leads to worst results (I-divergence is 2.71 for (d)) than if we suppose a Poisson noise (I-divergence is 0.69 for (c)).

We have also compared some other approaches as RL with Tikhonov-Miller (TM) regularization. The proposed method using TV gives better results than TM, both qualitatively and quantitatively: TM does not preserve the edges of the objects (they appear blurred), but TV does.

We also made the assumption for the noise to be white Gaussian noise instead of Poisson noise. We then used a gradient-descent type algorithm for deconvolution, regularized with TV. This time, the results are very good and close to those obtained with RL with TV regularization. Even if the noise statistics in confocal microscopy is proved to follow a Poisson law, we can say that the assumption of a Gaussian noise for confocal data is acceptable and leads to good results too.

5 Real data

We have also tested the proposed deconvolution algorithm on real data acquired with a confocal microscope. We have tested it on simple objects of known geometry: beads and spherical shells. We study the limit of the Poisson distribution hypothesis by deconvolving the same object acquired with different amounts of noise. It is more difficult to define a quantitative measure of the deconvolution results because we do not have any reference on the initial undegraded object (see section 3.5), but the geometrical data of the objects.

5.1 Microscope settings

The microscope and the objective we used were the same in all experiments. 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[10] 63x with numerical aperture $NA = 1.4$. The oil refractive index is 1.518 (23^o C). The acquisition software is Zeiss LSM 510 Meta which stores image stacks together with acquisition information in its file format. It is a kind of 3D TIFF format including many owner tags. We extract a 2D TIFF image sequence from this stack that we can read with standard libraries.

Here we define the noise level as the number of scanner passes for each line of the image acquisition: this is the "line average" function, in the microscope's acquisition software. A noise level of 2 means that the scanner has measured 2 times each line of the acquired image before taking the mean. Note that a noise level of 2 corresponds to a less noisy image than a noise level of 1.

5.2 Test objects

We are using several real specimens shown in fig. 11. We are working with small fluorescent beads[11] (with diameters 6 (a) and 15 μm (b)) and a spherical shell (c) of diameter (15 μm) and thickness between 0.5 to 0.7 μm according to Molecular Probes[12].

5.3 Deconvolution of real data

We present the deconvolution of the test objects using the proposed method, based on Richardson-Lucy (RL) algorithm with Total Variation (TV) regularization. The simulated point spread functions (PSF) used for the deconvolution of Fig. 12 and Fig. 14 are not represented here, because these are resampling of the one represented in fig. 2. Fig. 3 shows the best observation of the real confocal PSF, not taken at Nyquist rate, that have to be compared to Fig. 2.

[10] Apochromat is a chromatic aberration correction. From `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*".

[11] Fluorescent microspheres are FocalCheck^{TM} F-24634.

[12] Molecular Probes is the main manufacturer of fluorescence test objects such as microspheres, calibration grids, etc. `http://www.probes.com`

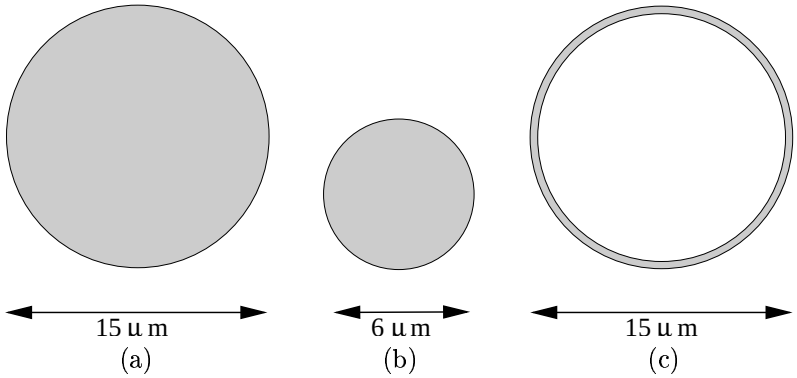
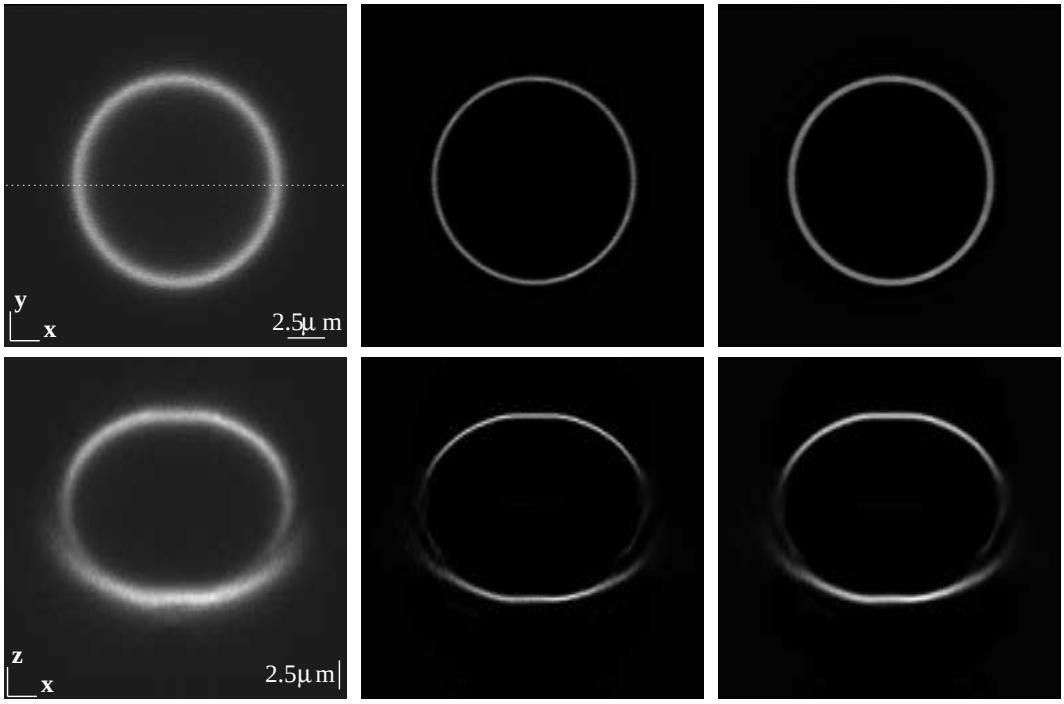


Figure 11: Fluorescent beads used in experiments. (a) and (b) are spherical beads with diameters 6 (a) and 15 μm (b); (c) is a spherical shell of diameter 15 μm stained with a fluorescent dye only on its surface (penetration depth between 500 to 700nm, according to manufacturer specifications).

Spherical shell We propose the deconvolution of a spherical shell of 15 μm. The acquisition parameters are the following:

- excitation wavelength: $\lambda_{ex} = 520\,\text{nm}$;
- emission wavelength: $\lambda_{em} = 488\,\text{nm}$;
- pinhole size: 1 Airy;
- line average: none;
- XY sampling: 89 nm per pixel (Nyquist);
- Z sampling: 230 nm per pixel (Nyquist);
- image volume: 256x256x128 pixels.

The results are shown in fig. 12: (a) are the raw data, (b) are the results of the deconvolution by standard RL, and (c) the results of the deconvolution using RL with TV. The noise amplification by RL occured after 486 iterations for (b) (I-divergence is 60.65), and the convergence of RL with TV (threshold of 10^{-4}) after 359 iterations (I-divergence is 60.63). There is no quantitative improvement between these two methods. Qualitatively, the results are nearly the same. The improvements we obtain from TV are better localization of the borders (see Fig. 13) and a better estimation of the shell thickness. As we have seen in section 5.2, the actual thickness of the shell is known to lie between 0.5 to 0.7 μm; in the restored image (a) we measure a thickness of 0.93μm which is too large, 0.26 in image (b) which is too small, and 0.40 in image (c), which also too small, but lies closer to the real value.



(a) (b) (c)

Figure 12: Deconvolution of a spherical shell image. First row represents the central XY image, and second line the central axial section of this stack corresponding to the dotted line. There is no resampling to obtain the same scale in X, Y and Z, only a resizing of the image. (a) represents the raw data, (b) the processed data using standard Richardson-Lucy, and (c) the processed data using RL and Total Variation regularization. The final values of I-divergence are 60.65 for standard RL and 60.63 for RL with TV.

Multiple object deconvolution We have tested the proposed deconvolution algorithm on a cluster of almost touching beads of diameter 6 μm (Fig. 11 (b)) to see if there is any interaction between objects during the deconvolution process. As we can see in fig. 14, the objects appear separate and the restoration intoduce no overlap.

For the deconvolution, we used a regularization parameter λ_{TV} of 0.005. The results of the deconvolution is not perfect, since there is still some blur in the three dimensions. However the residual blur may be due to a smooth gradient of the actual bead intensity. Nevertheless, there is no more noise in processed images (c), and there is no bead overlapping.

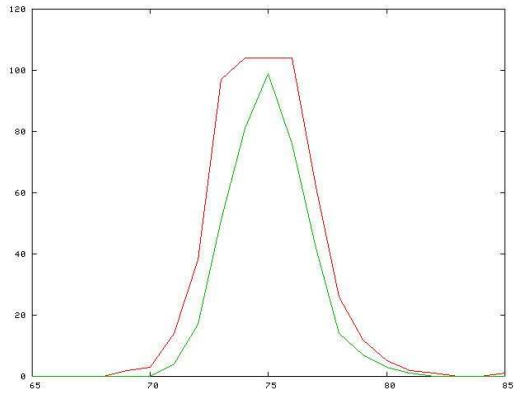


Figure 13: The TV reinforces the edges. Intensity profile of two XY images represented in fig. 12 (b) and (c), in the middle (cf. dotted line of Fig. 12 (a)). This is the right part of the ring. The outer (red) curve is the result of RL with TV and the inner (green) curve the result of standard RL. We can see that the TV regularization reinforces the edges of the shell, and that the restored shell thickness is larger. 1 pixel is 89nm.

5.4 The noise level and the deconvolution

We are working here on a single spherical shell, the same presented in section 5.3. We propose to study the influence of the noise on the deconvolution results by acquiring the same specimen four times under identical conditions except for a different line average (see section 5.1): no line average in fig. 12, a line average of 2 in fig. 15, of 4 in fig. 16 and of 8 in fig. 17. Poisson statistics is always verified, but we mathematically know that averaging Poisson distribution tends to a Gaussian distribution. We use the same parameters as previous experiments, except for the line averaging which is 2.

In each figure, from Fig. 15 to 17, the first row represents raw data, the second row processed data. As we have already seen, there is no line average in fig. 12 (the noise level is highest). On each figure, we can qualitatively see very good results with RL with TV: the blur and the noise are removed from each represented section, and the thickness of the shell is closer than the true value.

For each level of noise, the deconvolution creates a gap inside the image. This always corresponds to a transition zone (see YZ views on Fig. 12 to 17) between two non-connected circular arcs: one (top) is due to the fluorescence (image formation) but the other (bottom) circular arc is partially due to the reflexion of the laser on the mounting glass. But this gap can be easily explained knowing the properties of the TV. There are two zones of lower intensity at the extremities of the bottom circle arcs (YZ views of Fig. 12 to 17). As TV reiforces the edges, this intensity loss is considered as a "gap", and the borders are not connected together.

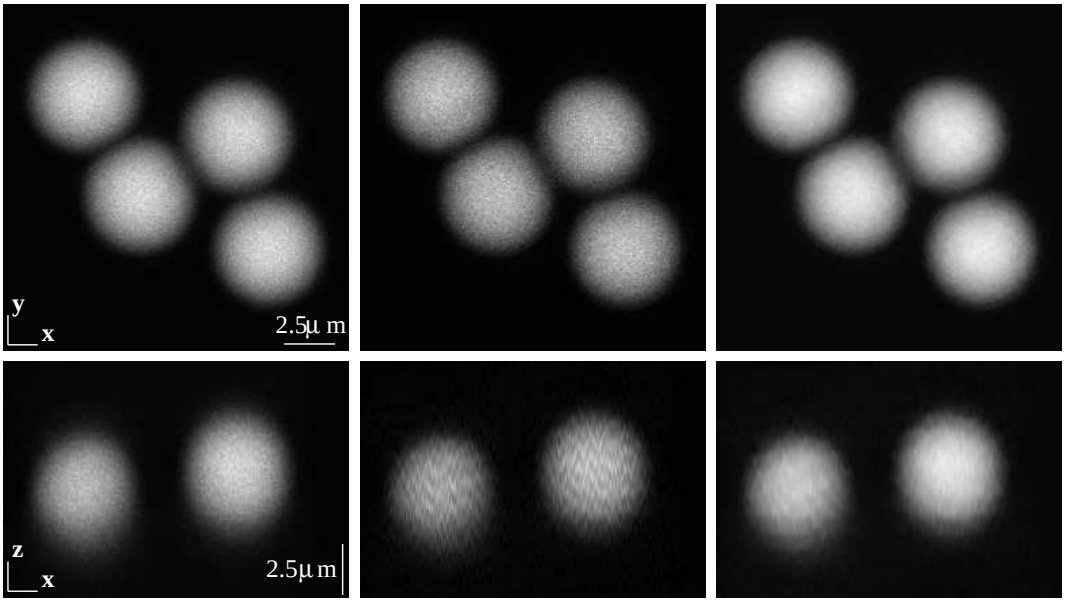
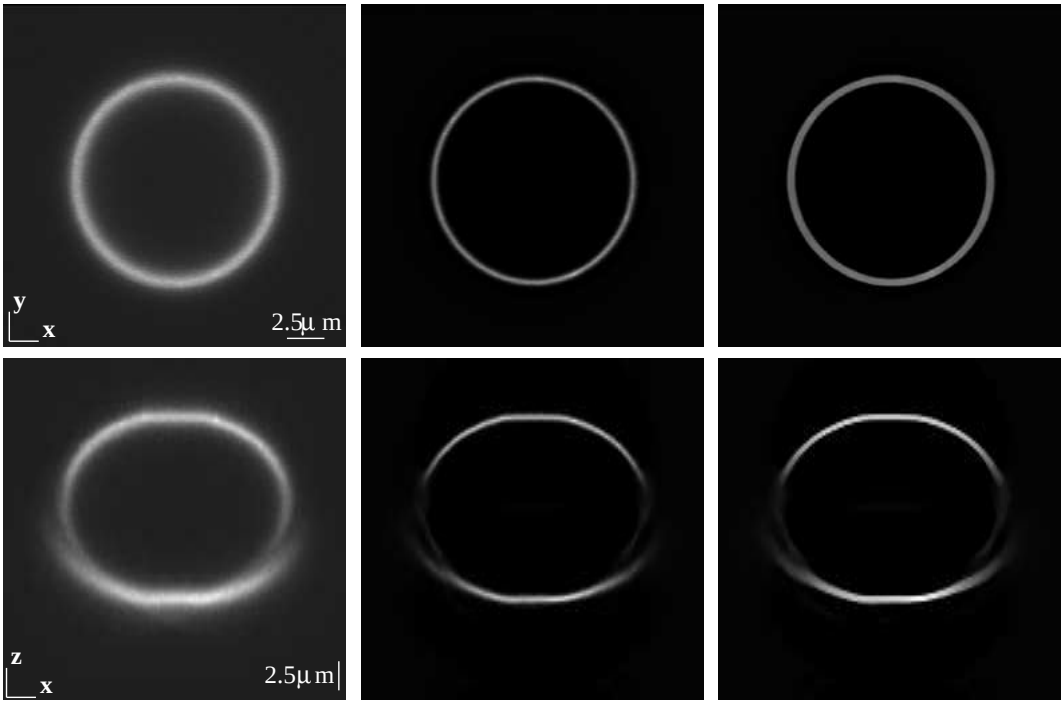


Figure 14: A cluster of 4 fluorescent beads of 6 μm. Top row shows XY views and bottom row shows YZ views. The stack is initially 256x256x128 with voxels of size 89x89x230nm. (a) raw data, (b) processed data using standard RL, and (c) processed data using RL with TV regularization ($\lambda_{TV} = 0.005$). (b) presents a high level of noise and the axial view shows some oscillation artefacts. In image (c), these degradations are no longer visible.

In contrast, we have a very good denoising and deblurring on all images, and we can notice that the thickness of restored shell decreases when the line average increases. This thickness decrease is most evident in YZ cuts (see Fig. 12 to 17, right column, YZ views). This is normal, as the noise becomes less important with increased averaging. But in fig. 17, for the most averaged acquisition, the thickness is too low compared to the values given by the manufacturer. For a line average of 8, as the image statistics is closer to a Gaussian statistics, the deconvolution algorithm based on Poisson statistics reaches its limits and should not be used.

5.5 Summary

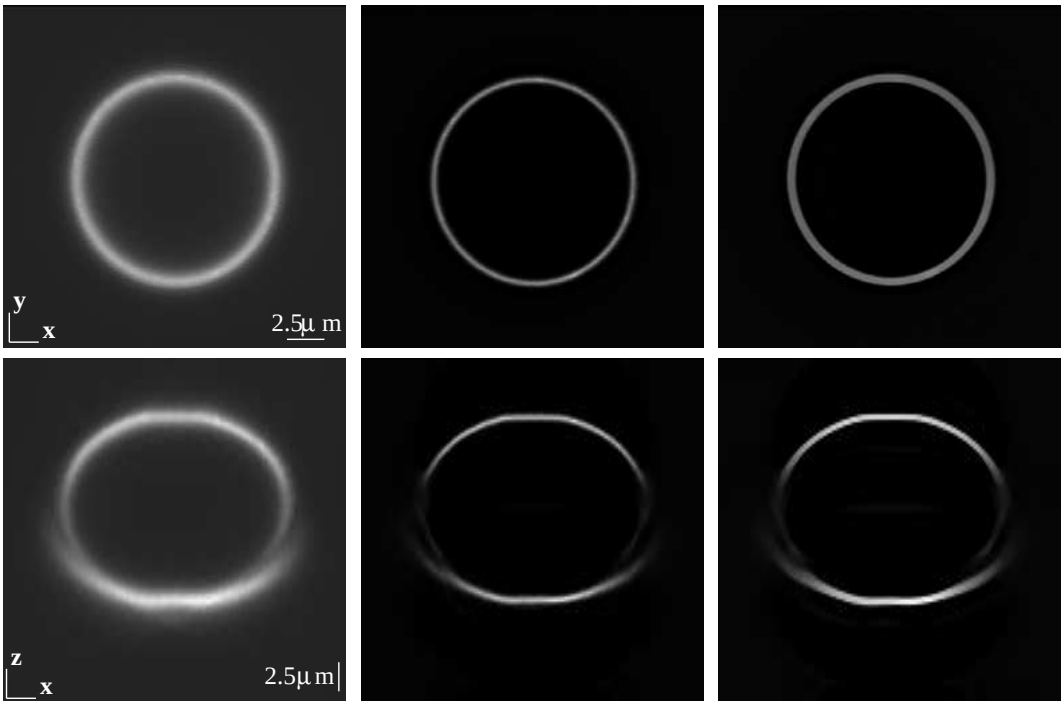
We have tested the deconvolution algorithm on several types of real objects: beads and spherical shells. The method gives good results both qualitatively and quantitatively: edges are sharper and noise is removed in the three dimensions. Nevertheless, we have noticed some limitations of the method. For images with moderate noise levels, the standard RL



(a) (b) (c)

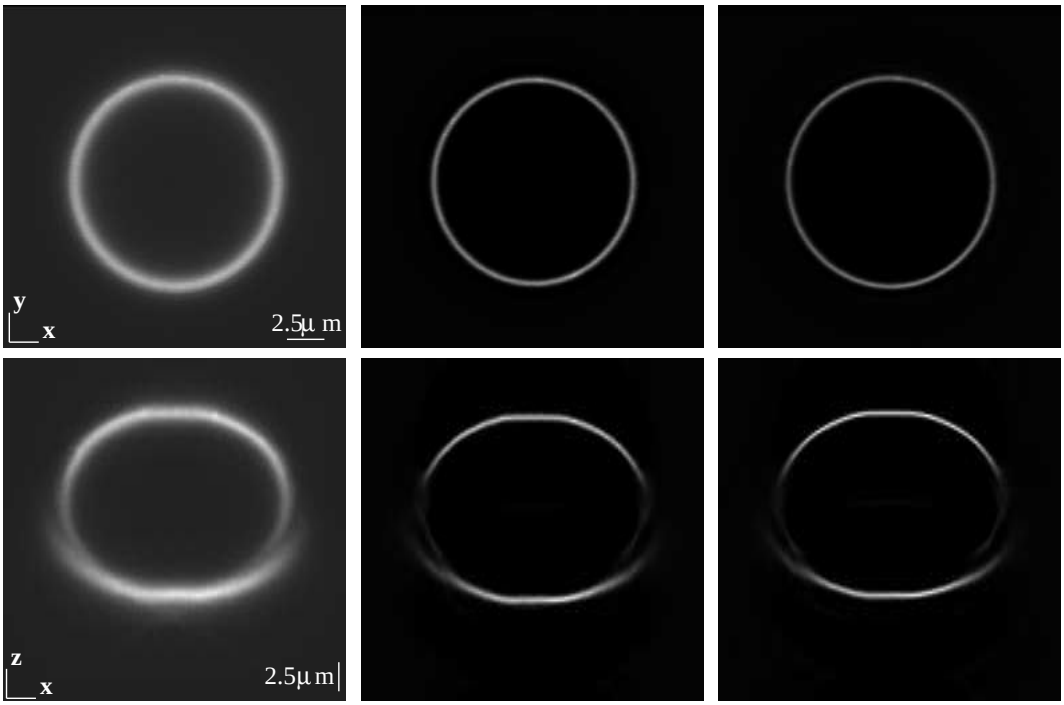
Figure 15: The spherical shell imaged with a line average of 2. Otherwise as in Fig. 12.

algorithm is sufficient to deconvolve the data (see, for instance, the images of the shell in fig. 12 to 17). The TV regularization effects are to make the edges sharper and more easy to detect for measuring the shell thickness. For objects with a high level of noise (see Fig. 14), standard RL is not sufficient for denoising.



(a) (b) (c)

Figure 16: The spherical shell imaged with a line average of 4. Otherwise as in Fig. 12.



(a) (b) (c)

Figure 17: The spherical shell imaged with a line average of 8. Otherwise as in Fig. 12.

6 Conclusion

6.1 Summary

In this research report, we have presented a new deconvolution approach for 3D confocal microscopy. It is based on the well-known Richardson-Lucy (RL) algorithm which is regularized using the Total Variation (TV) norm. RL with no regularization does not converge to an acceptable solution, since the solution is either non-convergent or dominated by the noise. In the literature, many authors have proposed to regularize this algorithm with different functionals (see section 2). Tikhonov-Miller functional is often used, but it over-smoothes the edges of objects in the image. We have proposed in this research report to use a new regularization term based on Total Variation, which is not quadratic and does not smooth edges (section 3). We have presented some results on simulated data (section 4) and on real data (section 5): the method gives very good results both qualitatively and quantitatively.

In [62, 65], van Kempen et al. have studied the influence of the background on the deconvolution. Actually, it speeds the convergence of the algorithm. They model the image formation process including the background. We do not take the background into account in the present paper because we make the assumption that estimating the background is a very hard regularization as it defines a strong support constraint. As we can see on results (see sections 4 and 5), the background is removed even with no early background estimation. We have studied the influence of the background in section 3.

The limitations of the proposed deconvolution algorithm with Total Variation regularization are known: for geometrical objects, the corners are rounded, and small details as well as texture are lost. This is an effect of the Total Variation, and this should be improved.

6.2 Future work

There are two different ways to improve the method: on the one hand, improving the functional that leads to the RL algorithm and the algorithm itself to avoid some numerical effects; on the other hand, the improvement concerns the point spread function (PSF) model, taking into account more details of the physics of the system. Another enhancement which could lead to a more user-friendly application would be an automatic parameter estimation (as done in 2D in [21, 26]), to have a more general method.

Improving the global image processing model As we already said, the TV regularization is not very adapted for fine structures such as actin filaments, which are often observed in biology and have diameters of around 8 nm and textures. The deconvolution of these kinds of objects will not lead to good results. In 2D, Malgouyres [33] proposes a hybrid model based on wavelet decomposition on the one hand, and a Total Variation based-denoising on the other. The wavelet decomposition permit to express and save the texture, while the algorithm based on TV is processing the rest of the image (only homogeneous part and noise). We plan to extend to 3D this idea by using both wavelets and TV-based deconvolution to restore real confocal images.

Another approach will be to extend the 2D non-iterative deconvolution model from Jalobeanu et al. [27]. This algorithm was designed for satelite image processing, but could be adapted to 3D confocal image restoration. This is a one-pass algorithm that first roughly deconvolves the image, with no regularization: the noise is strongly amplified, but the signal is still present. An adapted decomposition on a wavelet basis, followed by a thresholding (to remove noisy coefficients) leads to very good results in 2D.

Improving the image formation model On simulated data, we do not have to model a very precise PSF. Actually, as the same PSF is used for simulating the degradations over the image, and for the deconvolution, there is no risk of error for the PSF estimation. That is no longer the case when we are working with real data. Then we have to model accurately the PSF to have a suitable representation of the real PSF. To do this, we have to study very carefully the optical system and the specimen, to propose the most adapted model of PSF for the deconvolution.

We have seen (in section 2) that there are several possible models for a confocal microscope PSF. A direct measurement of the PSF is also possible but extremely sensitive to experimental conditions. However, this method has the advantage of taking into account the imperfections of the whole optical system. Any imperfection in the optical system influences the PSF: it is aberated.

When the PSF is aberated, it is not symmetric in the axial direction. These aberrations come from different sources: objective design (spherical aberration), bad alignment of the optical components, medium refractive index far from specimen refractive index (also known as refractive index mismatch), and many other.

Concerning the optical system of the microscope, spherical aberrations [17, 18, 20, 54] are the main optical aberrations. But some other aberrations come from the specimen and its preparation. Actually, the specimen itself and the mounting medium may induce many other aberrations if their refractive indexes are too different. This refractive index mismatch is discussed in [40] and [54].

In addition, to have a very precise PSF model, we have to take into account most of these aberration effects.

A Introduction to confocal microscopy

In this appendix, we briefly review the confocal microscopy principles. A very good reference can be found in [69] or on the website of Molecular Expression[13]. The confocal microscope was introduced by Minsky [36] in the 50^{th}. First we give some definitions of optical and deconvolution terminology. The second section explains how a confocal microscope works. We finally discuss the improvement of confocal microscopy relative to conventional microscopy.

Some definitions

Here we list terms that are used in this research report. Fig. 19 illustrates the main part of the microscope.

- `objective`: the objective (see Fig. 19) is an assembly of lenses which introduces the primary magnification of the system. It also defines the numerical aperture (NA) of the system;
- `numerical aperture`: the numerical aperture (NA) is related to the maximum angle of light ray α_{max} that can be collected by the objective. The NA is also proportional to the index of refraction n of the objective immersion liquid: $NA = n \sin \alpha_{max}$. The resolution of an objective is proportional to the NA, and the amount of light collected by an objective from an isotropic emitter (such as an excited fluorophore) is proportional to NA^2;
- `immersion medium`: an objective which is used without immersion medium is said to be dry, and then the NA value is at most 1. To increase this value, manufacturers have developped immersion objectives, often used with an oil[14] of refractive index of 1.515;
- `pinhole`: the pinhole is a small diaphragm (typically diameter around 1Airy Unit $\simeq$ 230 nm) which lets light emitted from the focal point pass through but rejects light from out-of-focus regions of the specimen. A completely opened diaphragm causes total loss of confocality of the microscope and the microscope is then completely equivalent to a widefield conventional fluorescence microscopy;
- `lateral directions`: these are the directions perpendicular to the observation axis, denoted by X and Y;
- `axial direction`: the direction along the observation axis, denoted by Z;
- `out-of-focus blur`: when imaging thick specimens, some parts of the specimens appear blurred. They are blurred because they are not exactly focused by the microscope. The PSF becomes wider with distance from the focus, although the integrated

[13] `http://www.microscopyu.com/articles/confocal/index.html`
[14] Biologists also use water-immersion objectives to observe directly live cells, tissues and embryos.

intensity of out-of-focus contributions is practically the same as in focus. The pinhole in the observation path reduces the out-of-focus blur;

- `low-photon imagery`: in low-photon imagery, only a few photons are detected per image pixel. The process of low photon counting is well represented by a Poisson process, and leads to multiplicative noise accumulation;
- `wavelength`: the wavelength λ of a light radiation (or for a photon) is related to the distance (measured in the direction of propagation) between two points in the same phase in consecutive cycles of a wave. The wavelength is related to the energy and the color of the radiation. It determines the minimal width of interference patterns, and therefore the resolution of an imaging system;
- `excitation wavelength`: the wavelength of the light that is generated by the laser, and used to excite fluorophores in the specimen;
- `emission wavelength`: the wavelength of the fluorescence light that is emitted by the fluorescent dye in the specimen;
- `Stokes shift`: the difference between the excitation and the emission wavelengths;
- `cut-off frequency`: the cut-off frequency of an optical system is the maximum spatial frequency that can enter the microscope objective. It is given by the physical limitations of the system, and the diffraction theory. For a confocal microscope, see [62, pp. 31-32];
- `lateral cut-off frequency`: the cut-off frequency $F_{e_{XY}}$ in the XY directions;
- `axial cut-off frequency`: the cut-off frequency F_{e_Z} for Z direction;
- `Nyquist lateral frequency`: the Nyquist lateral frequency $F_{max_{XY}}$ is the minimum sampling frequency that avoids spectral aliasing. The Nyquist lateral frequency is twice the lateral cut-off frequency $F_{max_{XY}} > 2\,F_{e_{XY}}$, and gives the smallest sampling frequency that allows to reconstruct the signal with no loss of information;
- `Nyquist axial frequency`: like Nyquist lateral frequency in XY, F_{max_Z} is twice the axial cut-off frequency to avoid aliasing when sampling the image in Z;
- `lateral sampling resolution`: d_{XY} is the sampling in XY at the lateral Nyquist frequency;
- `axial sampling resolution`: d_Z is the sampling in Z at the axial Nyquist frequency;
- `Nyquist lateral sampling`: the maximal sampling in XY that avoid aliasing;
- `Nyquist axial sampling`: the maximal sampling in Z with no aliasing;

Preparation of the specimen

Let us suppose that we want to observe a specimen using a confocal microscope. Before mounting the specimen on a glass slide, we choose one or more dyes to specifically label some parts of the specimen. These dyes are absorbed by the target parts of the specimen and fluoresce when they are illuminated by a beam of a specific wavelength. Fluorescence is explained in Fig. 18. When a photon is absorbed by the molecule, it is excited to a high electronic level. The molecule will quickly loose this energy in one or several ways, e.g. collision with other molecules or transition to other possible modes of vibration. Fluorescence occurs when the molecule returns to the electronic ground state, from the excited singlet state, by emission of a photon. This is not the only possible transition, but the only one leading to fluorescence.

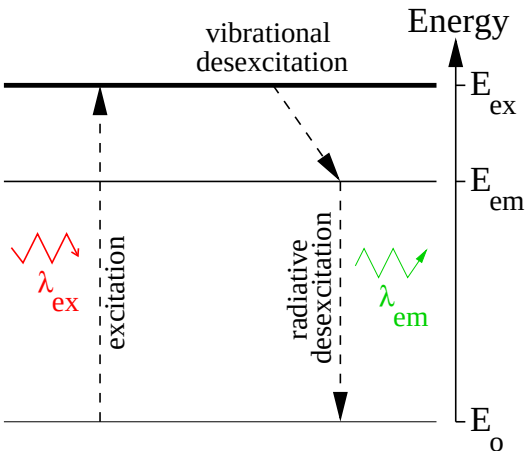


Figure 18: The process of fluorescence: if a photon of wavelength λ_{ex} is absorbed by a molecule of a dye, the molecule is excited. E_o is the ground state and E_{em} and E_{ex} two different excitation states. The vibrational energy of the excited state is usually different from the ground state; therefore the dye relaxes with p seconds losing some energy (non-radiative desexcitation) and then, after n seconds decay, emits a photon (fluorescence) or may also loose energy by creating a radical. During an image acquisition, this is known as the *bleaching* effect, which appears as a loss of intensity in time inside the specimen: more and more dye molecules are chemically transformed into other non fluorescent molecules. For more information, you can have further explications on the Jablonski energy diagram on `http://www.microscopy.fsu.edu/primer/java/jablonski/lightandcolor/`.

Confocal principle

Fig. 19 represents the image formation process in a confocal laser scanning microscope (CLSM). The excitation light is generated by a laser of wavelength λ_{ex}. The condenser lenses create a parallel beam which is mostly reflected by a dichroïc beam splitter to the objective lens. The objective lens focuses the light onto one point on the specimen. If this point contains a fluorescent dye, the dye will absorb the impinging light and become excited. After few nanoseconds, the fluorescent molecule returns to the ground state and reemits light isotropically. Due to the Stokes shift, the emitted light (i.e. the detected light) wavelength is longer. The intensity of the fluorescent emission is 10^{-5} or less of the intensity of the excited light, due to the absorption cross section, the fluorescent yield and the collection efficiency of the fluorescence detection. We can now understand why there is so little emitted light. The part of the light which goes through the objective passes through the beam splitter and the image formation lens focuses it to the pinhole, and behind to the photomultiplier. As the emission volume in the specimen is not infinitely small, some out-of-focus light is also collected by the objective. The pinhole (i.e. a very small aperture) between the imaging lens and the photomultiplier dramatically attenuate fluorescence which comes from out-of-focus.

Advantages of confocal over conventional microscopy

Reduction of the blur The comparison of imaging biological structures in conventional and confocal microscopy is discussed in [68]. The main advantage of confocal microscopy is the reduction of blur in the axial (Z) direction. It is also known as "background rejection" [46]. In Fig. 20, we represent in a geometrical way the principle of the confocal pinhole. For a conventional widefield microscope, there is no pinhole, and the contributions of blur in this case is very high compared to a confocal system.

In the lateral (XY) directions, the reduction of blur exists too. Looking at the formulation of the PSF (see Appendix B), the confocal PSF is approximatively the square of the widefield PSF. Its lateral size is thus considerably reduced[15]. These reductions of blur in all directions have the further advantage that it theoretically increases the resolution in the three directions.

Resolution increase The most common lateral resolution definition is given by the Rayleigh criterion. In any optical system, the image of one point is a diffraction pattern with maxima and minima. Considering the image of two source points of the same intensity, the Rayleigh criterion is:

[15] We cannot use the criterion of the first minima (i.e. the size of the PSF is defined between the first minima), because the widefield PSF and the confocal PSF have the same first minima. As there is an intensity loss, we have to define and estimate the new minimum that will define the resolution. This is really user dependant, but most of the specialists use a value of $\frac{1}{\sqrt{2}}$ or $\frac{1}{\sqrt{3}}$. If we want to compare this to a Gaussian (no minimum mathematically speaking), we may define the resolution limit as the full width at half maximum (FWHM); if the widefield PSF is a Gaussian, the confocal PSF is a Gaussian with a FWHM of $\frac{1}{\sqrt{2}}$ times the widefield one. In [10], the authors are discussing about the practical limits of the resolution in confocal microscopy.

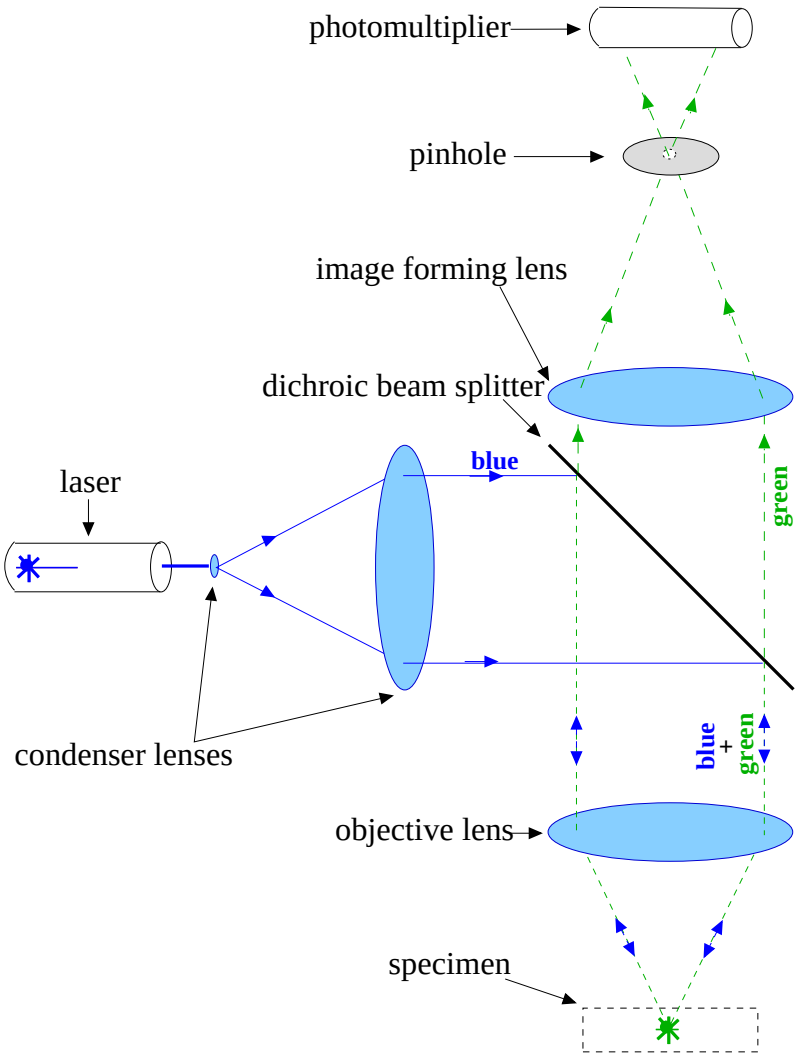


Figure 19: Confocal laser scanning microscope. The laser is the excitation light source. After spatially extending the laser beam, the light is partially reflected on a beam splitter. The reflected light arrives at the objective lens (bottom lens), which focuses it onto a specimen point. If it is fluorescent, it reemits some light. A part of this emitted light goes through the objective lens, and then the beam splitter. The image forming lens focuses the light on the photomultiplier (top), located just after the pinhole.

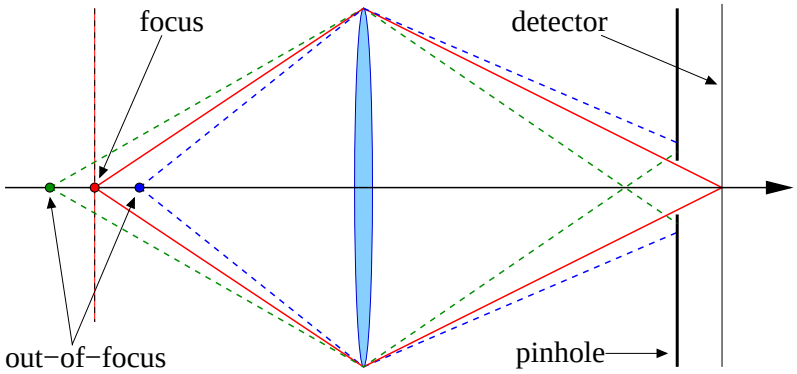


Figure 20: In a confocal system, the pinhole reduces the contributions of out-of-focus parts of the specimen: only focused regions are detected in full intensity.

> "Two components of equal intensity should be considered to be just resolved when the principal intensity maximum of one coincides with the first intensity minimum of other."[4, p. 371]

For a widefield microscope, the lateral resolution[16] is given by [4, p. 466]:

$$\gamma_{XY} = \frac{0.61\lambda_o}{NA} \tag{30}$$

λ_o is the excitation wavelength [61, p.4] in the vacuum, NA is the numerical aperture, and the factor 0.61 comes from the distance of the first minimum of an Airy function, the most common diffraction pattern when having a circular symmetry. The axial resolution is related to the depth-of-focus[17] of the microscope and is discussed in [52] in the widefield case. For a high aperture, and no paraxial approximation, the most common formulation is:

$$\gamma_Z = \frac{0.885\lambda_o}{n_o - \sqrt{n_o^2 - NA^2}} \tag{31}$$

n_o is the refractive index of the immersion oil of the objective. γ_Z is the distance over which the intensity is more than half the maximum. In confocal microscopy, resolution in XY and

[16] Many definitions of lateral resolution does exist. The main difference is the definition of the criterion concerning the higher maximum of the diffraction pattern, and the only change is the multiplicative constant: Rayleigh criterion has a constant of 0.61; sometime the definition could be the full width at half maximum (FWHM) of 0.51.

[17] Depth-of-focus is a subjective value depending on a definition. This explains why we can see different definitions for the depth-of-field in the litterature. In his paper, Sheppard [52] try to present the different formulations for a conventional microscope.

Z are divided by a factor 2 [61, p. 4]: the resolution in each directions is better. Thus, for a confocal microscope:

$$\begin{cases} \gamma_{XY}^{conf} = \frac{1}{\sqrt{2}}\gamma_{XY} \\ \gamma_{Z}^{conf} = \frac{1}{2}\gamma_{Z} \end{cases} \tag{32}$$

Sampling at a good rate

In the physical space, we have to deal with the system cut-off frequency that is $\rho_c^{inc} = 2\frac{NA}{\lambda_0}$ for a widefield system, or the system resolution given by the Rayleigh criterion, that is approximately the inverse $\gamma = \frac{0.61\lambda_0}{NA}$.

The Nyquist sampling theorem tells us that for a bandlimited system, the ideal sampling frequency should not be less than twice the cut-off frequency: $\rho_{sampling} \geq 2\rho_c^{inc}$. In the direct space, it is equivalent to set sampling distance (the "size" of one pixel) to less than

$$\Delta xy_{Nyquist}^{WF} = \frac{\lambda_0}{4\,NA} \tag{33}$$

For an immersion objective with 1.4 NA and a wavelength $\lambda_0 = 488$nm, we found $\Delta xy = 87$nm for a widefield system.

The fluorescent confocal case

In the fluorescent confocal case, the cut-off frequency is higher than the cut-off frequency for a conventional system: $\rho_c^{inc,\,conf} = 4\frac{NA}{\lambda_0}$. The Nyquist sampling distance is then

$$\Delta xy_{Nyquist}^{conf} = \frac{\lambda_0}{8\,NA} \tag{34}$$

For a confocal system with $NA = 1.4$ and $\lambda_0 = 488$nm, we find $\Delta xy_{Nyquist}^{conf} = 43.6$nm.

Actually, if we suppose the confocal to be perfect (infinitelly small pinhole), the PSF of the fluorescent system is [38, 53, 58, 59, 60] $PSF = \left|\hat{P}_{\lambda_{ex}}\right|^2 . \left|\hat{P}_{\lambda_{em}}\right|^2$. Supposing $\lambda_{ex} \simeq \lambda_{ex}$, that is equivalent to say that the OTF is the autocorrelation of the autocorrelation of the pupil function: the cut-off frequency is then four time the cut-off frequency of P.

The axial cut-off frequency

The axial cut-off frequency is related to the depth-of-field (DoF) of the microscope. The DoF is defined as the distance over which the intensity is more than half the maximum [52]:

$$\Delta z = 1.77\frac{\lambda_0}{4n\sin^2\frac{\alpha}{2}} = 1.77\frac{\lambda_0}{2n(1-\cos\alpha)} = 0.885\frac{\lambda_0}{\left(n-\sqrt{n^2-NA^2}\right)}$$

As for the xy resolution, we take the cut-off frequency in z to define the sampling distance:

$$w_c^{inc} = \frac{\left(n - \sqrt{n^2 - NA^2}\right)}{\lambda_0}$$

for a widefield system and $w_c^{inc,conf} = 2w_c^{inc}$ for a confocal system. The Nyquist distance in the axial direction is

$$\Delta z_{Nyquist}^{conf} = \frac{1}{4} \frac{\lambda_0}{\left(n - \sqrt{n^2 - NA^2}\right)}$$

For $\lambda_0 = 488$nm, $NA = 1.4$ and $n = 1.515$, we compute $\Delta z_{Nyquist}^{conf} = 130$nm.

Fixing a sampling value

Let us suppose we want to sample our image according to Nyquist's sampling rule. The physical size of one voxel becomes, in the confocal case, $\Delta xy_{Nyquist}^{conf}x\Delta z_{Nyquist}^{conf}$. We aquire an image volume of 256x256x32 which corresponds to a volume of $256\Delta xy_{Nyquist}^{conf}x256\Delta xy_{Nyquist}^{conf}$x-$32\Delta z_{Nyquist}^{conf}$ which correspond approximately to a volume in space of 11.1x11.1x4.2μm. In the frequency space, this is equivalent to say that the cut-off frequency is exactly on the (3D) boundaries of the image volume.

B The physics of the point spread function

The pupil function of an optical system

In an optical system, it is well-known that the image of one single point (written as the Dirac distribution $\delta(X,Y)$) is not really a point, but a diffraction pattern resulting from the limited apertures of the optical system and finite wavelength of light. Actually, only one finite-size aperture could be taken into account by considering only the exit pupil, defined as the most light limitting pupil inside all the optical system. Mathematically, the light wavefront in this pupil is written with a function called the pupil function P. We assume that we are working with incoherent monochromatic light[18], and the theory [4] shows that the image of one point is related with the squared modulus of the Fourier's transform of the pupil function.

Circular pupil function

Many shapes of pupil could be considered, but we only study here the case of a circular pupil of radius R giving the following pupil function:

$$P(X,\,Y) = \begin{cases} 1 \text{ if } \sqrt{X^2+Y^2} \leq R \\ 0 \quad \text{otherwise} \end{cases} \tag{35}$$

This pupil function is circularly symmetric, so we can redefine it as $P(r)$, $r = \sqrt{X^2+Y^2}$ being the radius from the origin. Now it is easy to calculate the Fourier's transform (FT) of this function, introducing $\rho = \sqrt{U^2+V^2}$, $\theta = \arctan \frac{Y}{X}$ and $\phi = \arctan \frac{V}{U}$:

$$\widehat{t}(\rho) = \int_0^R \int_0^{2\pi} e^{-2i\pi\rho r(\cos\theta\cos\phi+\sin\theta\sin\phi)} r\,dr\,d\theta \tag{36}$$

$$= \int_0^R r\,dr \int_0^{2\pi} e^{-2i\pi\rho r\,\cos(\theta-\phi)} d\theta \tag{37}$$

The zeroth order Bessel function is defined as $J_0(x) = \int_0^{2\pi} e^{-ix\,\cos(\theta-\phi)} d\theta$ then we can write:

[18] First assumption is monochromatic light and second assumption incoherent light.

$$\widehat{t}(\rho) = \int_0^R J_0(2\pi\rho r)\, r\, dr \tag{38}$$

$$= \frac{1}{2\pi\rho^2} \int_0^{2\pi\rho R} J_0(\omega)\, \omega\, d\omega \tag{39}$$

$$= \frac{1}{2\pi\rho^2} 2\pi\rho R J_1(2\pi\rho R) \tag{40}$$

$$= \pi R^2 \left(2\frac{J_1(2\pi\rho R)}{2\pi\rho R} \right) \tag{41}$$

$J_1(x)$ being the first order Bessel function. The first zero of this function occurs for $\rho R = 0.61$ (no unit). The incoherent PSF is proportional to the quantity $\left| 2\frac{J_1(2\pi\rho R)}{2\pi\rho R} \right|^2$, well-known as the "Airy pattern".

Circular pupil function in microscopy

In microscopy, we never access directly the radius of the pupil R, but microscope objectives always relate a dimensionless quantity, the numerical aperture (NA). Moreover, we know [4] that the first zero of the Airy pattern is given by the Rayleigh criterion $\gamma = \frac{0.61\lambda_0}{NA}$. This gives the resolution of the optical system, supposing that two points in the object space give two resolved points in the image space if the maximum of one's Airy pattern corresponds to the first zero of the other's. We deduce that the PSF is proportional to $\left| 2\frac{J_1(k_0 NA\, r)}{k_0 NA\, r} \right|^2$ with $k_0 = \frac{2\pi}{\lambda_0}$ is the wave vector. The relation between ρ and r is $\rho = \frac{NA}{\lambda_0 R}\, r$.

The 2D incoherent PSF and OTF

As we have already seen, the PSF of an optical system with monochromatic incoherent light is the squared modulus of the Airy pattern $PSF = |Airy|^2$. An Airy function is given by the FT of a circular pupil function, P. The PSF of the system is then $PSF = \left|\widehat{P}\right|^2$. The normalized PSF is:

$$h_{WF}^{inc}(r) = \frac{1}{N} \left| 2\frac{J_1(k_0 NA\, r)}{k_0 NA\, r} \right|^2 \text{ with } N = \int_{-\infty}^{+\infty} \left| 2\frac{J_1(k_0 NA\, r)}{k_0 NA\, r} \right|^2 dr \tag{42}$$

The optical transfer function (OTF) is the FT of the PSF. As $PSF = \left|\widehat{P}\right|^2$, the OTF is the autocorrelation of the pupil function $OTF = P * \overline{P}$ where $*$ denotes the convolution product and $\overline{P}$ the complex conjugate of P. P has a limited bandwith attained for $r = R$. Then,

the autocorrelation product (the OTF) has the double of this bandwith[19]. We can deduce the incoherent cut-off frequency of the system given for $r = 2R$, thus $\rho_c^{inc} = 2\frac{NA}{\lambda_0}$. To find an analytical form of the OTF, refer to [24, 55].

Numerical calculation of the PSF

Calculating the numerical values of the PSF knowing its analytical definition is very time consuming because of Bessel function computations. The best way to perform this, is to calculate first the OTF, and then the PSF with a Fourier transform. The autocorrelation product is calculated using the FT:

1. sample the defocused pupil function (frequencies space (U, V, Z));
2. perform its 2D FT (direct space (X, Y, Z));
3. take its squared modulus (PSF) (direct space (X, Y, Z));
4. perform its inverse FT (OTF) (frequency space (U, V, Z));

Please note that we are working in 2D: for each Z, we calculate the defocused pupil function, then perform a 2D FT, etc. To have a 3D PSF, we pile each 2D plane in the order of the Z value. The 3D OTF is then obtained by applying a 3D FFT on this 3D PSF.

The 3D incoherent PSF

In the widefield case, the 3D OTF for wavelength λ is defined by:

$$OTF_\lambda(U, V, W) = \left(P * \overline{P}\right)(U, V, W) \tag{43}$$

As a definition, the 3D PSF is the FT of the OTF:

$$PSF_\lambda(X, Y, Z) = Fourier_{3D}\left[OTF_\lambda(U, V, W)\right] \tag{44}$$

As the light passes 2 times through the objective, with two different wavelengths, the ideal confocal OTF is defined as:

$$OTF_{Conf}(U, V, W) = \left(p_{\lambda_{ex}} * \overline{P}_{\lambda_{ex}}\right) * \left(P_{\lambda_{em}} * \overline{P}_{\lambda_{em}}\right)(U, V, W) \tag{45}$$

and its PSF:

$$PSF_{Conf}(U, V, W) = \left(PSF_{\lambda_{ex}}\right)\left(PSF_{\lambda_{em}}\right)(X, Y, Z) \tag{46}$$

[19]It is easy to demonstrate with a scheme: the autocorrelation product is proportional to the common area of the P spectrum and its sliding complex conjugate spectrum. It is zero when these two spectrum stop to overlap, which is when the zero frequencies are distant from one time the bandwith.

λ_{ex} being the excitation wavelength and λ_{em} the emission wavelength. If the Stokes shift (see Appendix A) is zero, the confocal PSF is the square of the widefield PSF:

$$PSF_{Conf}(U,V,W) = (PSF_{\lambda})^2 (X,Y,Z) \tag{47}$$

which is a usual approximation.

The size of the pinhole

Until now, we have supposed that the pinhole (which limits the out-of-focus contributions of the light) was a perfect one. Its mathematical definition is then given by a 2D Dirac function $A(X,Y) = \delta(X,Y)$. Using the definition $r^2 = X^2 + Y^2$, we have $A(r) = \delta(r)$. On real confocal systems, the pinhole has a non-zero physical size, because it can not be infinitely small. Its size is more often given in Airy units (AU) than in nanometers[20]. The advantage of the AU is that it takes the light wavelength into account. A typical size of a usable pinhole is $1\,AU$, since it allows more than 70 % of the in-focus light to reach the PMT detector. The final PSF, containing both the contributions of the excitation light, the emission light and the pinhole, is given by:

$$PSF(r,Z) = \left\{ A(r) * \left| \widehat{P}_{\lambda_{ex}}(r,Z) \right|^2 \right\} . \left| \widehat{P}_{\lambda_{em}}(r,Z) \right|^2 \tag{48}$$

If the pinhole is perfect, we find the FT of Eq. 45, and if the pinhole is very large, the first part of Eq. 48 vanishes and we have the PSF of a widefield system. The pinhole of a radius of R' is given by:

$$A(r) = \begin{cases} 1 \text{ if } r \leq R' \\ 0 \text{ elsewhere} \end{cases} \tag{49}$$

The final computation of the PSF in our model

We describe here how we compute the 3D PSF of a confocal microscope in our implementation. We just suppose that we know the needed physical values of our system:

- for the objective, its magnification (M) and its numerical aperture (NA);
- for the pinhole, its diameter;
- for the light, both its emission (λ_{em}) and excitation (λ_{ex}) wavelengths;

[20] scale: $1\,AU = 0.61 \frac{\lambda}{NA}$ nanometers, if λ is in nanometers.

The first part $A(r) * \left|\widehat{P}_{\lambda_{ex}}(r,Z)\right|^2$

1. Calculate the pupil function $P_z(r) = P(r,Z)$ for a given defocus Z;
2. Doing the same for all Z in the range of the axial cut-off frequency;
3. Performing a 2D FT to obtain $\widehat{P}_Z(r)$;
4. Take its square modulus;
5. Piling up numerically all the $\left|\widehat{P}_Z(r)\right|^2$;
6. Convolve it with $A(r)$.

The computation require many 2D FT (depending on the Z sampling) for step 3, and two 3D FFT for step 6..

The second part $\left|\widehat{P}_{\lambda_{em}}(r,Z)\right|$

1. Calculate the pupil function $P_Z(r) = P(r,Z)$ for a given defocus Z;
2. Doing the same for all Z in the range of the axial cut-off frequency;
3. Performing a 2D FT to obtain $\widehat{P}_Z(r)$;
4. Take its square modulus;
5. Piling up numerically all the $\left|\widehat{P}_Z(r)\right|^2$;
6. Multiply it by the first part $A(r) * \left|\widehat{P}_{\lambda_{ex}}(r,Z)\right|^2$

The computation require many 2D FT depending on the Z sampling for step 3, and no 3D FFT.

C The deconvolution algorithms in detail

In this appendix, we present the algorithms that are used in this research report in more detail. We also give the derivations that lead to these algorithms. First we present some algorithms derived under the hypothesis of a Poisson noise: the standard Richardson-Lucy (RL) algorithm and then, RL with different regularizations such as Tikhonov-Miller and Total Variation. At the end of this appendix, we propose an additive gradient-descent algorithm based under the assumption of a Gaussian noise. We present this algorithm with and without Total Variation regularization.

Image formation

We assume the following image formation model: $i(x) = (o * h)(X)$. i is the image, o the object and h the impulse function of the system, known as the point spread function (PSF). The variable x is $2-$ or $3-$dimensions and is (X, Y) or (X, Y, Z). This equation is written under the assumption of no noise. If we have Poisson noise (such in confocal microscope imaging), we have to deal with $i = \wp\,(o * h)$, $\wp$ being the Poisson distribution. We remind that with a Bayesian approach, we can write:

$$p(o \mid i) = p(i \mid o).\frac{p(o)}{p(i)} \tag{50}$$

where $p(i \mid o)$ is the likelihood probability, $p(o \mid i)$ the a posteriori probability, and $p(o)$ a prior model on the object.

In the presence of a Poisson process

In a deconvolution problem, if we assume that the image statistics is described by a Poisson process, the likelihood probability could be expressed as [30]:

$$p(i \mid o) = \prod_x \frac{\left[(h * o)\,(x)\right]^{i(x)}\, e^{-(h*o)(x)}}{i(x)\,!} \tag{51}$$

where $i = \wp\,(h * o)$, $\wp$ being a Poisson process. One way to solve this problem is to maximize equation 51, which is equivalent to minimize:

$$-\log p(i \mid o) = \int_x \left[(h * o)\,(x) - i(x).\log\,(h * o)\,(x) + \log\,(i(x)\,!)\right] dx \tag{52}$$

where the term $\log\,(i(x)\,!)$ in Eq. 52 is a constant relatively to o. We can thus define the new functional to minimize as:

$$J_1(o) = \int_x \left[(h * o)\,(x) - i(x).\log\,(h * o)\,(x)\right] dx \tag{53}$$

Early minimization

Let us suppose that the functional $s(x)$ is a 2D or 3D image, and that the variable x could be (X, Y) or (X, Y, Z); ρ is a parameter. Considering a small perturbation ρs of o, Eq. 53 becomes:

$$J_1(o+\rho s) = \int_x \left[(h * (o+\rho s))(x) - i(x).\log\left(h * (o+\rho s)\right)(x)\right] dx \tag{54}$$

Omitting the (x) dependence in the notations, we rewrite Eq. 54 as:

$$J_1(o+\rho s) = \int_x \left[(h*o) + \rho\,(h*s) - i.\log\left[(h*o)\left(1+\rho\frac{(h*s)}{(h*o)}\right)\right]\right] dx \tag{55}$$

$$\simeq \int_x \left[(h*o) + \rho\,(h*s) - i.\log(h*o) - \rho i.\frac{(h*s)}{(h*o)}\right] dx \tag{56}$$

$$= J_1(o) + \rho \int_x \left[(h*s) - i.\frac{(h*s)}{(h*o)}\right] dx \tag{57}$$

Calculating the solution of $\nabla J_1 = 0$

Defining the scalar product $\langle f,\, g\rangle = \langle g,\, f\rangle = \int_x (f.g)(x)dx$, we have the following equality for the derivative of J_1:

$$\left\langle \frac{\partial J_1}{\partial o}(o),\, \rho s \right\rangle = \lim_{\rho \to o^+} \frac{J_1(o+\rho s) - J_1(o)}{\rho} \tag{58}$$

Expressing the second term of Eq. 57 as a scalar product leads to:

$$\int_x \left[(h*s) - i.\frac{(h*s)}{(h*o)}\right] dx = \langle 1,\, s*h\rangle - \left\langle \frac{i}{(h*o)},\, h*s \right\rangle \tag{59}$$

$$= \langle h^*,\, s\rangle - \left\langle \frac{i}{(h*o)} * h^*,\, s \right\rangle \tag{60}$$

$$= \int_x s.\left[h^* - h^* * \left(i.\frac{1}{(h*o)}\right)\right] dx \tag{61}$$

Here we have used the notation h^* for the adjoint of the operator h, using the fact that $\langle g,\, f*s\rangle = \langle g*s^*,\, f\rangle$ (not shown here). For the PSF h, we may also write that $h^*(x) = h(-x)$. We are now able to express $\nabla_o J_1 = \frac{\partial J_1}{\partial o}$ as:

$$\nabla J_1(o) = h(-x) * \left[1 - \frac{i(x)}{(h*o)(x)}\right] \tag{62}$$

We can now minimize Eq. 62, i.e. solve $\nabla J_1(o) = 0$:

$$\int_x h(-x)dx - h(-x) * \frac{i(x)}{(h * o)(x)} = 0 \tag{63}$$

We assume that the PSF $h(x)$ is normalized to 1, then $\int_x h(-x)dx = 1$. This is a well used assumption in optics. We have:

$$h(-x) * \frac{i(x)}{(h * o)(x)} = 1 \tag{64}$$

The multiplicative algorithm: Richardson-Lucy

By solving Eq. 64 in a multiplicative way, we can derive the following algorithm, by assuming that at convergence, the ratio $\frac{o_{k+1}(x)}{o_k(x)}$ is 1:

$$o_{k+1}(x) = o_k(x) \left[h(-x) * \frac{i(x)}{(h * o_k)(x)} \right] \tag{65}$$

This is the well-known Richardson-Lucy (RL) algorithm. RL algorithm has the interesting property of non-negativity: if the first estimate (actually $o_0(x)$) is non-negative, none of the further estimate will be negative. Nevertheless, for noisy images, RL does not converge to a suitable solution in the general case, as it amplifies the noise after several iteration; one must stop iterations before this happens. Note that the algorithm usually converges slowly to the suitable solution at first, and wanders slowly from this solution after a while. We can choose a stop criterion based on the absolute number of iterations, or based on the difference between two estimations o_k and o_{k+1} for instance, such as $s = \frac{\sum_x o_{k+1}(x) - o_k(x)}{\sum_x o_k(x)}$. Another problem with RL is its non-convergence with no regularization and the denominator which may have some zero values. We choose to set to 0 every value for which the denominator of $o_{k+1}(x)$ is too small.

The additive form of the algorithm

Still under the assumption of a Poisson process, we can also use the additive version of the RL algorithm given by Eq. 65:

$$o_{k+1} = o_k + \delta t \left(1 - h(-x) * \frac{i(x)}{(h * o_k)(x)} \right) \tag{66}$$

This is an additive gradient-descent algorithm based on a Poisson process; the parameter δt regulates the step of the descent, i.e. the step between two iterations.

Regularizations

We have seen that the standard RL does not always converge to a suitable solution. That is because we have not introduced any information on the object. If we propose a prior model on the object, we regularize the solution. We will now maximize the a posteriori probability instead of maximizing the likelihood probability.

The 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 $p_{TM}(o) = N_{TM} e^{-\lambda_{TM} \int_x |\nabla o(x)|^2}$. λ_{TM} is the regularization parameter for TM regularization term, and N_{TM} is the normalization constant for the probability. To take into account the object model, we have to maximize the a posteriori probability $p(o \mid i)$. We have to minimize the functional:

$$J_2(o) = \int_x \left[(h * o)(x) - i(x).\log\left(h * o\right)(x)\right] dx + \lambda_{TM} \int_x |\nabla o(x)|^2 \, dx \tag{67}$$

The derivative of the first term gives the same result as Eq. 62, and the minimization of the regularization term (that we call J_{TM}) gives:

$$\begin{aligned} J_{TM}(o + \rho s) &= \int |\nabla (o + \rho s)|^2 && (68)\\ &\simeq \int |\nabla o|^2 + 2\rho \int (\nabla s \nabla o) && (69)\\ &= J_{TM}(o) + 2\rho \langle \nabla s, \nabla o \rangle && (70)\\ &= J_{TM}(o) + 2\rho \langle s, \nabla^* \nabla o \rangle && (71)\\ &= J_{TM}(o) - 2\rho \langle s, \triangle o \rangle && (72) \end{aligned}$$

where $\triangle = -\nabla^* \nabla$ stands for the Laplacian ($\triangle o = \frac{\partial^2 o}{\partial x^2} + \frac{\partial^2 o}{\partial y^2} + \frac{\partial^2 o}{\partial z^2}$). We can apply for J_{TM} the same equation as Eq. 58 to finally obtain the functional minimized relative to o:

$$\int_x h(-x)dx - h(-x) * \frac{i(x)}{(h * o)(x)} - 2\lambda_{TM} \triangle o(x) dx = 0 \tag{73}$$

To solve this equation by using a multiplicative approach (see Eq. 65), we obtain a regularized version of the RL algorithm:

$$o_{k+1}(\mathbf{x}) = \left\{ \left[\frac{i(\mathbf{x})}{(o_k * h)(\mathbf{x})} \right] * h(-\mathbf{x}) \right\} . \frac{o_k(\mathbf{x})}{1 + 2\lambda_{TM} \triangle o_k(\mathbf{x})} \tag{74}$$

The Total Variation regularization The Total Variation (TV) is a regularization method introduced by Rudin et al. [45] on the gradient of the image. The continuous and the discrete forms are described in Appendix D. When used as a prior model for the object, we have to minimize the following functional:

$$\mathrm{Min}_o \left[J_1\left(o(x)\right) + \lambda \int_x \left|\nabla o(x)\right| dx \right] \tag{75}$$

We first replace the Total Variation member by its smooth approximation, by adding a *small* constant β^2 that avoid some problems near the origin (see [15]). Expanding the Total Variation term $J_{TV}(o) = \int_x \sqrt{\left|\nabla o(x)\right|^2 + \beta^2} dx$ of a small perturbation $o + \rho s$ of o yields:

$$J_{TV}(o+\rho s) = \int_x \sqrt{\left(\nabla o(x)\right)^2 + \rho^2 \left(\nabla s\right)^2 + 2\rho \nabla o.\nabla s + \beta^2} dx \tag{76}$$

$$\simeq \int_x \sqrt{\left(\nabla o(x)\right)^2 + \beta^2} . \sqrt{1 + 2\rho \frac{\nabla o.\nabla s}{\left(\nabla o(x)\right)^2 + \beta^2}} dx \tag{77}$$

$$\simeq J_{TV}(o) + \rho \int_x \frac{\nabla o}{\sqrt{\left(\nabla o(x)\right)^2 + \beta^2}} \nabla s \, dx \tag{78}$$

$$\simeq J_{TV}(o) + \rho \int_x \frac{\nabla o}{|\nabla o|} \nabla s \, dx \tag{79}$$

$$= J_{TV}(o) + \rho \left\langle \frac{\nabla o}{|\nabla o|}, \nabla s \right\rangle \tag{80}$$

$$= J_{TV}(o) + \rho \left\langle \nabla^* \frac{\nabla o}{|\nabla o|}, s \right\rangle \tag{81}$$

$$= J_{TV}(o) - \rho \int_x \mathrm{div}\left(\frac{\nabla o}{|\nabla o|} \right) s \, dx \tag{82}$$

Where we used the notation $\nabla^* = -\mathrm{div}$. The minimization of $J_1(o)$ is the same as Eq. 62 and gives, for the first terms:

$$\int_x h(-x) dx - h(-x) * \frac{i(x)}{\left(h * o\right)(x)} - \lambda \mathrm{div}\left(\frac{\nabla o}{|\nabla o|} \right) = 0 \tag{83}$$

λ being the regularization parameter that could be adaptive [21] or a constant [14]. If we derive a multiplicative algorithm in the same way as we obtained Eq. 65, we have RL regularized with TV:

$$o_{k+1}(x) = \frac{o_k(x)}{1 - \lambda \mathrm{div}\left(\frac{\nabla o_k}{|\nabla o_k|} \right)} \left[h(-x) * \frac{i(x)}{\left(h * o_k\right)(x)} \right] \tag{84}$$

The greatest problem with the multiplicative form (Eq. 84) is that the regularization term is in the denominator. The whole denominator could be zero or negative, violating the major property of the standard RL. That forces us to choose a small value for λ around 10^{-3}. Nevertheless, this version of the algorithm is prooved to be convergent.
To avoid these problems and still under the assumption of a Poisson process, we can propose the additive form of the algorithm. The additive form regularized with TV from Eq. 66 is:

$$o_{k+1}(x) = o_k(x) + \delta t \left[-1 + \lambda \mathrm{div}\left(\frac{\nabla o_k}{|\nabla o_k|} \right) + h(-x) * \frac{i(x)}{(h * o_k)(x)} \right] \tag{85}$$

Properties and limitations of Total Variation

Briefly, the main property of the Total Variation regularization is that it regularizes the images while preserving the borders [33]. The linear Tikhonov-Miller regularization, for instance, has not this property and tends to smooth the borders in the image. The limitations of Total Variation based methods is that the textures (actualy, the small structures close to the noise) are not well restored (staircase effect) and that some corners are enlarged [33].

Under the assumption of a Gaussian distribution

If we suppose that the image statistics follows a Gaussian distribution instead of a Poisson one, we can model the image formation as:

$$i(x) = o(x) * h(x) + n_G(x) \tag{86}$$

where i is the observed image, o the initial object blurred by the optical system PSF h, and corrupted with an additive white Gaussian noise n_G. Under this hypothesis, the likelihood becomes [25, pp.32]:

$$p(i \mid o) = \frac{1}{(2\pi)^{\frac{N_x N_y N_z}{2}} \sigma} e^{-\frac{\|i-(o*h)\|^2}{2\sigma^2}} \tag{87}$$

where σ is the standard deviation of the noise, $N_{x,y,z}$ are the size of the image in the x, y and z directions. The maximization of this likelihood leads to the following functional to minimize:

$$J_{+_o}(o) = A + B \|i - (o * h)\|^2 \tag{88}$$

A and B being two positive constants with no need to be expressed. The resolution of the minimization of J_{+_o} in an additive way leads to a gradient-descent based algorithm:

$$o_{k+1}(\mathbf{x}) = o_k(\mathbf{x}) + \alpha \left[(h^* * i)(\mathbf{x}) - (h^* * h) * o_k(\mathbf{x}) \right] \tag{89}$$

The index k is the iteration index, the variable $\mathbf{x}$ represents the 3D discrete directions along x, y and z, and the exposant $*$ denotes the adjoint operator (here the adjoint of h is h^*, and $h^*(\mathbf{x}) = h(-\mathbf{x})$).

If we introduce the Total Variation (TV) as a prior model on the object, we now have to maximize the a posteriori probability $p(o \mid i)$. The new functional to be minimized relatively to o becomes:

$$J_+(o) = J_{+_o}(o) + \lambda \sum_{\mathbf{x}} |\nabla o(\mathbf{x})| \tag{90}$$

λ is the regularization parameter like in Eq. 22. we get:

$$o_{k+1}(\mathbf{x}) = o_k(\mathbf{x}) + \alpha \left[(h^* * i)(\mathbf{x}) - (h^* * h) * o_k(\mathbf{x}) + \frac{\lambda}{\alpha} div \left(\frac{\nabla o_k}{|\nabla o_k|} \right) \right] \tag{91}$$

We express the regularization parameter as $\frac{\lambda}{\alpha}$ instead of another λ_+. To do a comparison, we use the same values of λ as with RL algorithm with a TV regularization. This gradient-based algorithm is additive, and is more stable than a multiplicative algorithm.

D The 3D total variation routine

The 3D total variation is used to regularize Richardson-Lucy algorithm presented in Appendix C. In a continuous form, the total varition of any 3D functional f is given as:

$$J_{TV}(f) = \lambda \int_x |\nabla f(x)| \, dx \tag{92}$$

where λ is the TV regularization parameter. Using L^2 norm, we have:

$$J_{TV}(f) = \lambda \int\int\int \sqrt{f_x^2 + f_y^2 + f_z^2} \, dx \, dy \, dz \tag{93}$$

Its derivation gives:

$$\frac{\partial J_{TV}}{\partial f}(f) = -\lambda \operatorname{div}\left(\frac{\nabla f}{|\nabla f|}\right) \tag{94}$$

The discrete form of $\operatorname{div}\left(\frac{\nabla f}{|\nabla f|}\right)$ for a stable numerical scheme is the following [45]:

$$\begin{aligned}\operatorname{div}\left(\frac{\nabla f}{|\nabla f|}\right) = {} & \frac{1}{h_x}\Delta_-^x \frac{\Delta_+^x f_{ijk}}{\sqrt{\left(\Delta_+^x f_{ijk}\right)^2 + m\left(\Delta_+^y f_{ijk}, \Delta_-^y f_{ijk}\right)^2 + m\left(\Delta_+^z f_{ijk}, \Delta_-^z f_{ijk}\right)^2}} \\ & + \frac{1}{h_y}\Delta_-^y \frac{\Delta_+^y f_{ijk}}{\sqrt{\left(\Delta_+^y f_{ijk}\right)^2 + m\left(\Delta_+^z f_{ijk}, \Delta_-^z f_{ijk}\right)^2 + m\left(\Delta_+^x f_{ijk}, \Delta_-^x f_{ijk}\right)^2}} \\ & + \frac{1}{h_z}\Delta_-^z \frac{\Delta_+^z f_{ijk}}{\sqrt{\left(\Delta_+^z f_{ijk}\right)^2 + m\left(\Delta_+^x f_{ijk}, \Delta_-^x f_{ijk}\right)^2 + m\left(\Delta_+^y f_{ijk}, \Delta_-^y f_{ijk}\right)^2}}\end{aligned} \tag{95}$$

defining the derivation as:

$$\begin{array}{ll} \Delta_+^x f_{ijk} = \frac{1}{h_x}\left(f_{(i+1)jk} - f_{ijk}\right) & \Delta_-^x f_{ijk} = \frac{1}{h_x}\left(f_{ijk} - f_{(i-1)jk}\right) \\ \Delta_+^y f_{ijk} = \frac{1}{h_y}\left(f_{i(j+1)k} - f_{ijk}\right) & \Delta_-^y f_{ijk} = \frac{1}{h_y}\left(f_{ijk} - f_{i(j-1)k}\right) \\ \Delta_+^z f_{ijk} = \frac{1}{h_z}\left(f_{ij(k+1)} - f_{ijk}\right) & \Delta_-^z f_{ijk} = \frac{1}{h_z}\left(f_{ijk} - f_{ij(k-1)}\right) \end{array} \tag{96}$$

and the function $m(a,b)$ as:

$$m(a,b) = \frac{\operatorname{sign} a + \operatorname{sign} b}{2} \min(|a|, |b|) \tag{97}$$

The function $\operatorname{sign}(a)$ returns 1 for $x > 0$, -1 for $x < 0$ and 0 for $x = 0$. One advantage of $m(a,b)$ is that this function equals zero if a and b have the opposite sign. Thus, if $\Delta_+^{x,y,z} f_{ijk}$ and $\Delta_-^{x,y,z} f_{ijk}$ have not the same sign, we choose $\Delta_\pm^{x,y,z} f_{ijk} = 0$. We use the following boundary limits $\frac{\partial f}{\partial x_i} = 0$ ($x_1 = x$, $x_2 = y$, $x_3 = z$) for a continuous space. For a discrete functional, this leads to:

$$\begin{array}{ccc} f_{0jk} = f_{1jk} & f_{i0k} = f_{i1k} & f_{ij0} = f_{ij1} \\ f_{(N_x+1)jk} = f_{(N_x)jk} & f_{i(N_y+1)k} = f_{i(N_y)k} & f_{ij(N_z+1)} = f_{ij(N_z)} \end{array} \tag{98}$$

This is exactly the algorithm that has been implemented in `C++`.

E Preliminary results on biological samples

Figs. 21 and 22 show views extracted from a 3D sequence, imaging a real biological specimen with some very fine structures. Fig. 21 shows a regenerating mammalian skin with an alveolus network, while Fig. 22 shows the same specimen with some filaments. We use different values for the regularisation parameter from $\lambda_{TV} = 0.001$ to 0.01. We do not show YZ views because of the low number of samples we have (30). We clearly see the disappearance of some details if the regularisation parameter is too high, and the presence of noise if it is too small.

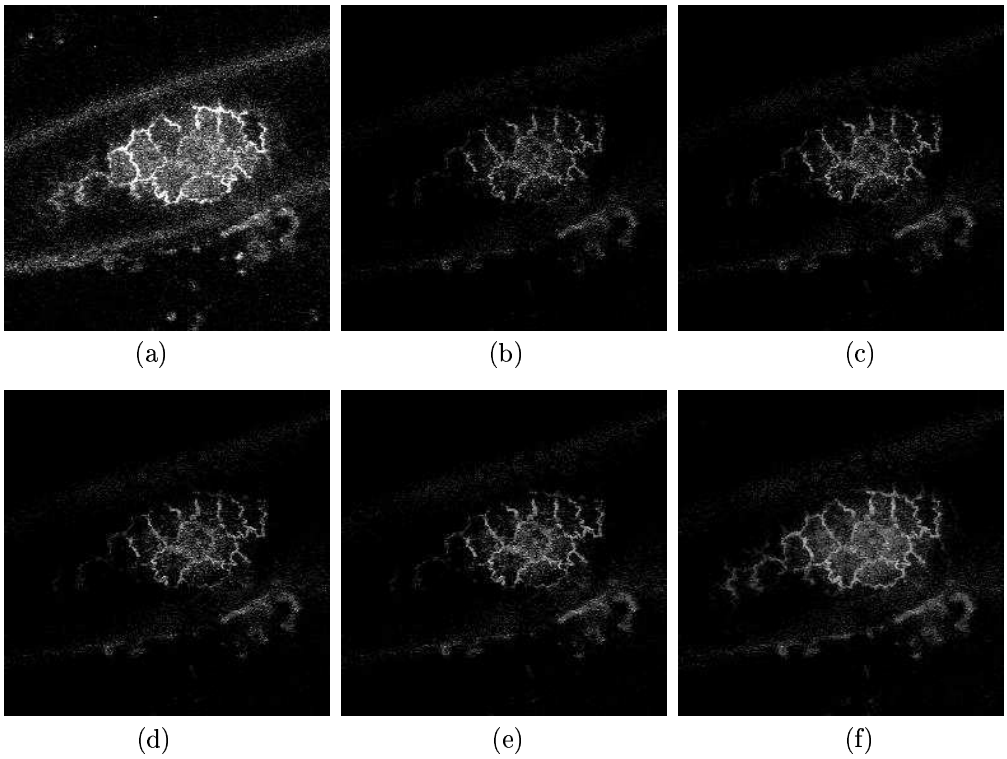


Figure 21: A sample of a regenerating mamalian skin showing alveolus structures. It is extracted from a sequence of 30 images. The microscope is a Zeiss Axiovert 200 with an objective of 40x, $NA = 1.3$. (a) raw images; (b) standard RL deconvolution; (c)-(f) RL with TV with different values of the regularisation parameter λ_{TV}: (c) 0.001; (d) 0.0025; (e) 0.005; (f) 0.01. The raw image (a) is very noisy. The noise is still present for standard RL deconvolution (b) and more and more disappears while λ_{TV} is increasing ((c) to (f)). Qualitatively, (f) gives the best results. (© UMR 6543 CNRS/Laboratoire J.-A. Dieudonné)

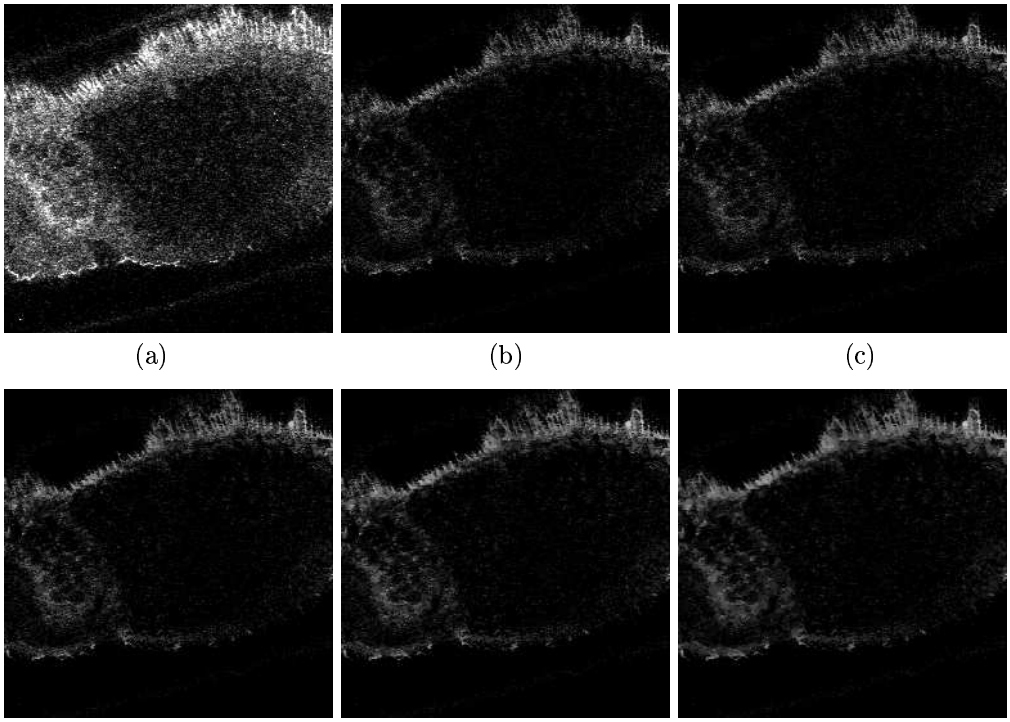


Figure 22: RL with TV on real fine structures (same specimen as in Fig. 21). (a) raw images; (b) standard RL deconvolution; (c)-(f) RL with TV with different values of the regularisation parameter λ_{TV}: (c) 0.001; (d) 0.0025; (e) 0.005; (f) 0.01. We can see that the filaments are strongly degraded by the noise (top of figure (a)). The deconvolution with standard RL (b) or with RL with TV with a small parameter (c) leaves some noise. With RL and TV with intermediate parameters (d)-(e), the noise is strongly removed and the filaments are visible, but for strong value (f) of λ_{TV}, the noise is removed but the filaments are smoothed together. (© UMR 6543 CNRS/Laboratoire J.-A. Dieudonné)

References

[1] D.A. Agard, Y. Hiraoka, P. Shaw, and J.W. Sedat. Fluorescence microscopy in three dimensions. *Methods Cell Biol.*, 30:353–377, 1989.

[2] D.A. Agard and J.W. Sedat. Three-dimensional architecture of a polytene nucleus. *Nature*, 302:676–681, April 1983.

[3] M.R. Arnison and C.J.R. Sheppard. A 3D vectorial optical transfer function suitable for arbitrary pupil functions. *Optics Communications*, 211:53–63, 2002.

[4] M. Born and E. Wolf. *Principles of Optics.* Cambridge University Press, 7th (expanded) edition, 1999.

[5] K. Carlsson. Confocal imaging for 3-D digital microscopy. *Applied Optics*, 26(16):3232–3238, August 1987.

[6] K.R. Castleman. *Digital Image Processing.* Prentice Hall, 1996.

[7] P. Charbonnier. *Reconstruction d'image : régularisation avec prise en compte des discontinuités.* PhD thesis, Université de Nice-Sophia Antipolis, 1994.

[8] J.-A. Conchello and J.G. McNally. Fast regularization technique for expectation maximization algorithm for optical sectioning microscopy. In *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing III*, pages 199–208, 1996.

[9] J.-A. Conchello and Q. Yu. Parametric blind deconvolution of fluorescence microscopy images: preliminary results. In *Proceedings of the 1996 IS&T/SPIE symposium on electronic imaging: Science and technology*, pages 164–174, April 1996.

[10] G. Cox and C.J.R. Sheppard. Practical limits of resolution in confocal and non-linear microscopy. *Microscopy Research and Technique*, 263:18–22, January 2004.

[11] I. Csiszár. Why least squares and maximum entropy? *The Annals of Statistics*, 19:2032–2066, 1991.

[12] J. Boutet de Monvel, S. Le Calvez, and M. Ulfendahl. Image restoration for confocal microscopy: Improving the limits of deconvolution, with application to the visualization of the mammalian hearing organ. *Biophysical Journal*, 80:2455–2470, May 2001.

[13] J. Boutet de Monvel, E. Scarfone, S. Le Calvez, and M. Ulfendahl. Image-adaptive deconvolution for three-dimensional deep biological imaging. *Biophysical Journal*, 85:3991–4001, December 2003.

[14] N. Dey, L. Blanc-Féraud, C. Zimmer, Z. Kam, J.-C. Olivo-Marin, and J. Zerubia. A deconvolution method for confocal microscopy with total variation regularization. In *Proceedings of ISBI'2004*, April 2004.

[15] D.C. Dobson and C.R. Vogel. Convergence of an iterative method for total variation denoising. *SIAM Journal of Numerical Analysis*, 34(5):1779–1791, October 1997.

[16] S. Durand, F. Malgouyres, and B. Rougé. Image deblurring, spectrum interpolation and application to satellite imaging. *Control Optimisation and Calculus of Variation*, 5:445–477, 2000.

[17] L. Baker Editor. *Selected papers on : Effects of Aberrations in Optical Imaging.* SPIE Milestone Series, Volume MS 59, Optical Engeneering Press, 1992.

[18] V.N. Mahajan Editor. *Selected papers on : Effects of Aberrations in Optical Imaging.* SPIE Milestone Series, Volume MS 74, Optical Engeneering Press, 1994.

[19] A.R. FitzGerrel, E.R. Dowski, and W. T. Cathey. Defocus transfer function for circularly symmetric pupils. *Applied Optics*, 36(23):5796–5804, August 1997.

[20] S.F. Gibson and F. Lanni. Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy. *Journal of Optical Society of America A*, 8(10):1601–1613, October 1991.

[21] G. Gilboa, N. Sochen, and Y.Y. Zeevi. Texture preserving variational denoising using an adaptive fidelity term. In *Proceedings of VLSM'2003*, October 2003.

[22] J. W. Goodman. *Introduction to Fourier Optics.* McGraw-Hill Book Company, 2nd edition, 1996.

[23] P.J. Green. On use of the EM algorithm for penalized likelihood estimation. *Journal of Royal Statist. Soc. B*, 52(3):443–452, 1990.

[24] H.H. Hopkins. The frequency response of a defocused optical system. *Proceedings of the Royal Society of London Serie A*, 231:91–103, February 1955.

[25] A.K. Jain. *Fundamentals of digital image processing.* Prentice-Hall, 1993.

[26] A. Jalobeanu, L. Blanc-Féraud, and J. Zerubia. Hyperparameter estimation for satellite image restoration by a MCMC maximum likelihood method. *Pattern Recognition*, 35(2):341–352, 2002.

[27] A. Jalobeanu, L. Blanc-Féraud, and J. Zerubia. Satellite image deblurring using complex wavelet packets. *IJCV*, 51(3):205–217, 2003.

[28] C. Kervrann and A. Trubuil. An adaptive window approach for poisson noise reduction and structure preserving in confocal microscopy, April 2004.

[29] S. Kimura and C. Munakata. Calculation of a three-dimensional optical transfer function for a confocal scanning fluorescent microscope. *Journal of Optical Society of America A*, 6(7):1015–1019, July 1989.

[30] K. Lange. Convergence of EM image reconstruction algorithm with Gibbs smoothing. *IEEE Transactions on Medical Imaging*, MI-9(4):439–446, December 1990.

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

[32] F. Malgouyres. A framework for image deblurring using wavelet packet basis. *Applied and Computational HArmonic Analysis*, 12(3):309–331, 2002.

[33] F. Malgouyres. Mathematical analysis of a model which combines total variation and wavelet for image restoration. *Journal of Information Processes*, 2:1–10, 2002.

[34] J. Markham and J.-A. Conchello. Parametric blind deconvolution: a robust method for the simultaneous estimation of image and blur. *Journal of Optical Society of America A*, 16(10):2377–2391, October 1999.

[35] J. Markham and J.-A. Conchello. Fast maximum-likelihood image-restoration algorithms for three-dimensional fluorescence microscopy. *Journal of Optical Society of America A*, 18(5):1062–1071, 2001.

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

[37] S. Osher, A. Solé, and L. Vese. Image decomposition, image restoration, and texture modeling using total variation minimization and the h^{-1} norm. In *Proceedings of ICIP'2003*, September 2003.

[38] J.B. Pawley. *Handbook of Biological Confocal Microscopy*. Plenum Press, New York, 2nd edition, 1996.

[39] J.B. Pawley. *Sources of Noise in Three-Dimensional Microscopical Data Sets*. Academic Press, Inc., 1996.

[40] J.B. Pawley. Limitations on optical sectionning in live-cell confocal microscopy. *Scanning*, 24(5):241–246, September 2002.

[41] M. Persson, D. Bone, and H. Elmqvist. Total variation norm for three-dimensional iterative reconstruction in limited view angle tomography. *Physics in Medecine and Biology*, 46:853–866, March 2001.

[42] J. Polzehl and V. Spokoiny. Adaptive weigths smoothing with application to image restoration. *Journal of Royal Statistical Society B*, 62(2):335–354, 2000.

[43] C. Preza, M.I. Miller, L.J. Jr. Thomas, and J.G. McNally. Regularized linear method for reconstruction of three-dimensional microscopic objects from optical sections. *Journal of Optical Society of America A*, 9(2):219–228, February 1992.

[44] W. H. Richardson. Bayesian-based iterative method of image restoration. *JOSA*, 62:55–59, 1972.

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

[46] D.R. Sandison, D.W. Piston, and W.W. Webb. *Background rejection and optimization of signal to noise in confocal microscopy.* Academic Press, San Diego, 1994.

[47] L.H. Schaefer, D. Schuster, and H. Herz. Generalized approach for accelerated maximum likelihood based image restoration applied to three-dimensional fluorescence microscopy. *Journal of Microscopy*, pages 99–107, January 2001.

[48] A. Schönle and S.W. Hell. Calculation of vectorial three-dimensional transfer functions in large-angle focusing systems. *Journal of Optical Society of America A*, 19(10):2121–2126, October 2002.

[49] Z. Shao, O. Baumann, and A.P. Somlyo. Axial resolution of confocal microscopes with parallel-beam detection. *Journal of Microscopy*, 164(1):13–19, October 1991.

[50] C.J.R. Sheppard. The spatial frequency cut-off in three-dimensional imaging. *Optik*, 72(4):131–133, 1986.

[51] C.J.R. Sheppard. The spatial frequency cut-off in three-dimensional imaging II. *Optik*, 74(3):128–129, 1986.

[52] C.J.R. Sheppard. Depth of field in optical microscopy. *J. of Microscopy*, 149:73–75, January 1988.

[53] C.J.R. Sheppard and C.J. Cogswell. Three-dimensional image formation in confocal microscopy. *Journal of Microscopy*, 159:179–194, August 1990.

[54] C.J.R. Sheppard and P. Török. Effects of specimen refractive index on confocal imaging. *Journal of Microscopy*, pages 366–374, March 1997.

[55] P.A. Stokseth. Properties of a defocused optical system. *Journal of Optical Society of America*, 59(10):1314–1321, October 1969.

[56] A.N. Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Soviet Math. Dokl.*, 57(4):1035–1038, 1963.

[57] A.N. Tikhonov and V.Y. Arsenin. *Solution of ill-posed problems.* Winston-Wiley, New York, 1977.

[58] H.T.M. van der Voort. *Three Dimensional Image Formation and Processing in Confocal Microscopy.* PhD thesis, Amsterdam University, November 1989.

[59] H.T.M. van der Voort and G.J. Brakenhoff. 3-D image formation in high-aperture fluorescence confocal microscopy: a numerical analysis. *Journal of Microscopy*, pages 43–54, April 1990.

[60] H.T.M. van der Voort and K.C. Strasters. Restoration of confocal images for quantitative image analysis. *Journal of Microscopy*, pages 165–181, May 1995.

[61] H. van der Voort et al. *SVI Image Restoration Recipes.* 2002.

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

[63] G.M.P. van Kempen, H.T.M. van der Voort, and L.J. van Vliet. A quantitative comparison of two restoration methods as applied to confocal microscopy. In *Proceedings of ASCI'96, 2nd Annual Conference of the Advanced School for Computing and Imaging*, pages 196–201, June 1996.

[64] G.M.P. van Kempen and L.J. van Vliet. The influence of the background estimation on the superresolution properties of non-linear image restoration algorithms. In *Proceedings of the SPIE Conference on Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing IV*, volume 3605, pages 179–189, January 1999.

[65] G.M.P. van Kempen and L.J. van Vliet. Background estimation in non linear image restoration. *Journal of Optical Society of America A*, 17(3):425–433, March 2000.

[66] 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*, 198:63–75, April 2000.

[67] 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, March 1997.

[68] J.G. White, W.B. Amos, and M. Fordham. An evaluation of confocal versus conventional imaging of biological structures by fluorescence light microscopy. *Journal of Cell Biology*, 105:41–48, July 1987.

[69] S. Wilhelm, B. Gröbler, M. Gluch, and H. Heinz. *Confocal Laser Scanning Microscopy: Optical Image Formation - Electronic Signal Processing.* Carl Zeiss Advanced Imaging Microscopy, 2001.

[70] R. Willett and R.D. Nowak. Platelets for multiscale analysis in photon-limited imaging. In *Proceedings of ICIP'2002*, 2002.

[71] R. Willett and R.D. Nowak. Platelets: a multiscale approach for recovering edges and surfaces in photon-limited medical imaging. *IEEE Trans. on Medical Imaging*, 22(3):332–350, March 2003.

[72] L.T. Young, R. Zagers, L.J. van Vliet, J.Mullikin, F. Boddeke, and H. Netten. Depth-of-focus in microscopy. In *Scandinavian Conference on Image Analysis*, volume 1, pages 493–498, 1993.

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

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

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