A "LEARN 2D, APPLY 3D" METHOD FOR 3D DECONVOLUTION MICROSCOPY.

Ferréol Soulez

Centre de Recherche Astrophysique de Lyon
CNRS-UMR 5574 – Université Lyon 1 – ENS Lyon, Université de Lyon, France.

ABSTRACT

This paper presents a 3D deconvolution method for fluorescence microscopy that reached the first place at the "the 3D Deconvolution Microscopy Challenge" held during ISBI 2013. It uses sparse coding algorithm to learn 2D "high resolution" features that will be used as a prior to enhance the resolution along depth axis. This is a three steps method: (i) deconvolution step with total variation regularization, (ii) denoising of the deconvolved image using learned sparse coding, (iii) deconvolution using denoised image as quadratic prior. Its effectiveness is illustrated on both synthetic and real data.

1. INTRODUCTION

1.1. Motivation

Thanks to the widespread of fluorescent labeling techniques, fluorescent microscopy is one of the most used imaging modality in biology and epifluorescence (a.k.a. wide field) microscopes are the simplest set up for such observations. It consists on imaging at its emission wavelength a cellular structure marked by fluorescent dye excited by uniform illumination. On the resulting 2D image, structures are more or less defocalized according to their distance to the focal plane. Moving this focal plane through the sample produces a 3D representation of the object. However, its characteristics both in terms of resolution and noise are often too coarse to fit in constraints of current biological research. Furthermore the resolution is not isotropic and the actual resolution along the depth axis is far more coarser than the resolution along lateral dimensions. Beside more advanced microscope designs such as the confocal or multiphotons microscopes, one can use only computational methods numerically invert the blurring process and enhance epifluorescence micrographs as demonstrated by Agard and Sedat [1]. Since, many deconvolution algorithms were proposed (see [2, 3] for reviews) and these *"deconvolution microscopy"* methods lead to various commercial and open-source software [4, 5] but still suffer of several drawbacks: computational and memory costs, restoration artifacts, non-quantitative results... In the 2013, the first edition of the *"3D Deconvolution Microscopy Challenge"* was organized to stimulate the community to look for novel, global and practical approaches to this problem. I present in this paper the 3D deconvolution method that was awarded in this challenge.

1.2. Notations

In this document I use the following notations:

- $\boldsymbol{y} \in \mathbb{R}^N$ is the observed data;
- $N = N_x \times N_y \times N_z$ is the size of the observed data;
- $\boldsymbol{g} \in \mathbb{R}^K$ is the processed data;
- $K = K_x \times K_y \times K_z$ is the size of the processed data;
- $k = 1, \dots, K$ is the voxels index;
- $\mathbf{H} \in \mathbb{R}^{K \times K}$ is the forward operator;
- $\boldsymbol{w} \in \mathbb{R}^K$ is the vector of statistical weights;
- $\boldsymbol{u}^+ \in \mathbb{R}^K$ is the solution of the first step;
- $\boldsymbol{u}^{++} \in \mathbb{R}^K$ is the solution of the second step;
- $\boldsymbol{u}^{+++} \in \mathbb{R}^K$ is the solution of the last step;
- $s \in \{1, 2, 3\}$ is the 2D subspace index corresponding to the sections $\{(x, y), (x, z), (y, z)\}$ respectively;
- $m = 1, \dots, M^{[\mathrm{s}]}$ is the pixel index in patches of size $M^{[\mathrm{s}]}$;
- $\boldsymbol{p} \in \mathbb{R}^{(M^{[1]}+M^{[2]}+M^{[3]}) \times K}$ is the patch vector with $p_{m,k,s}$ the value of the m^{th} pixel of the patch of the orientation s of the slice centered on pixel k;
- $\mathcal{P}^{[\mathrm{s}]}$ is the set of all overlapping patches laying in the subspaces s;
- $\ell = 1, \dots, L$ is the atom index in dictionary containing L atoms;
- $\mathbf{D} \in \mathbb{R}^{L \times (M^{[1]}+M^{[2]}+M^{[3]})}$ is the dictionary;
- $\mathcal{D}$ is the set of dictionary with L atoms of size M with unit ℓ_2 norm;
- $\boldsymbol{\alpha} \in \mathbb{R}^{L \times K \times 3}$ is the vector of coefficients;
- $\mathbf{R} \in \mathbb{R}^{K \times K \times M \times 3}$ is a reconstruction operator;

2. ALGORITHM

2.1. Principle

As mentioned in the introduction, the resolution along the depth axis is coarser (at best 2 times) than the resolution along lateral dimensions. In most deconvolution method, the resolution of the deconvolved data is still coarser in depth than laterally. The key idea of this work is to extract "high resolution" features from lateral (x, y) sections and use them as a prior to enhance the lower resolution along depth axis. It means that I implicitly rely on that object structures are statistically isotropic (no preferred orientation).

This "copy - paste" method is achieved using sparse coding on a learned dictionary [6, 7]. In sparse coding, every patch of the denoised image is a linear combination of few dictionary elements. In learned sparse coding, this dictionary is learned on a training set or directly using noisy image patches. The main contribution of this work is to learn the dictionary only using 2D patches extracted from lateral (x, y) sections of the 3D image of the previously deconvolved image. This learned dictionary is then used to denoise every 2D images of every sections, namely (x, y), (x, z) and (y, z), which are then averaged.

The algorithm we propose here has three steps:

work supported by projects MiTiV and POLCA funded by the French National Research Agency (ANR DEFI 09-EMER-008-01 and ANR-10-BLAN-0511).

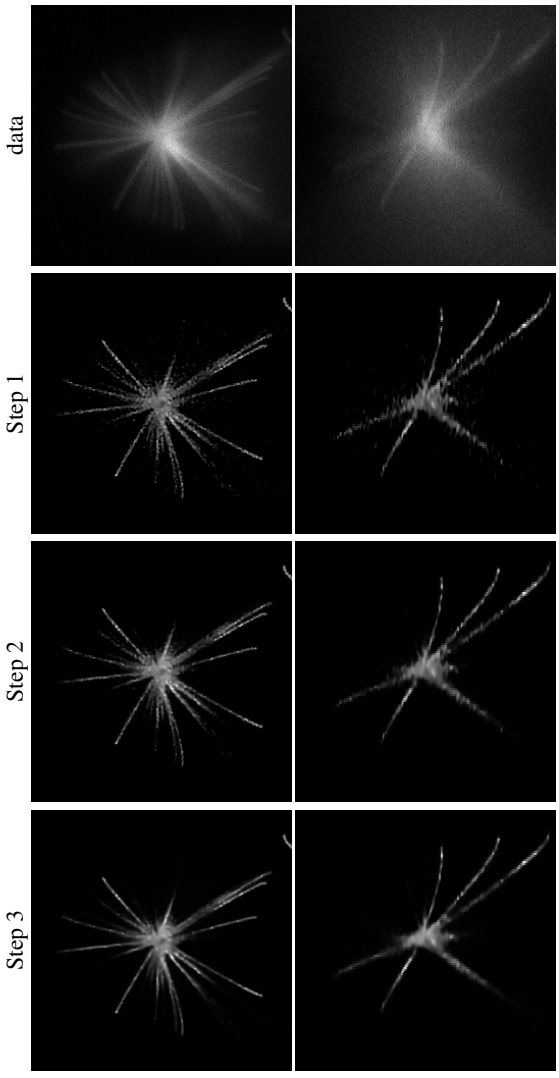


Fig. 1. Results of the three steps of our method on parts of the green channel of the first *"3D Deconvolution Microscopy Challenge"* data set.

1. deconvolution using Total Variation regularization,
2. denoising of the deconvolved image using learned sparse coding,
3. deconvolution using denoised image as quadratic prior.

The results of these three steps are illustrated on figure (1) on a zoom of the green channel (with filaments) of the deconvolution challenge data set.

2.2. Step 1: Total variation deconvolution

This first step is exactly identical to the deconvolution step described in [8]. The *maximum a posteriori* estimates $\boldsymbol{u}^+$ obtained by minimizing a cost function under a non-negativity constraint:

$$\boldsymbol{u}^+ = \underset{\boldsymbol{u}\geq 0}{\arg\min} \left(\mathcal{J}_{\mathsf{data}}(\boldsymbol{u}) + \mu_1 \, \mathcal{J}_{\mathsf{TV}}(\boldsymbol{u}) \right) , \tag{1}$$

with $\mathcal{J}_{\mathsf{data}}(\boldsymbol{u})$ a data fitting term, $\mathcal{J}_{\mathsf{TV}}(\boldsymbol{u})$ a regularization term and μ_1 an hyper-parameter.

2.2.1. Data-fitting term

Under uncorrelated Gaussian noise assumption, the data fitting term writes:

$$\mathcal{J}_{\mathsf{data}}(\boldsymbol{u}) = \sum_k w_k \left([\mathbf{H} \cdot \boldsymbol{u}]_k - g_k \right)^2 , \tag{2}$$

where w_k is the inverse of the noise variance at voxel k. This model can cope with non-stationary noise and can be used to express confidence on measurements on each voxel of the data. Thus it can deal with unmeasured voxels (due to saturation, borders…) by setting $w_k = 0$ for such voxels. With such a formulation, it is possible to take rigorously the borders into account. In order to apply $\mathbf{H}$ using FFT-based computations, most works suppose that $\mathbf{H}$ is circulant. Doing so, to keep a rigorous direct model without the unrealistic periodic assumption, both data and estimated object must be extended. This field extension is mandatory for two reasons: (i) to prevent spatial aliasing and (ii) to estimate parts of the scene outside of the field of view that contributes to the measurements inside the instrument field of view [9]. For that reason, as the PSF has the same size as the data, we extend the reconstructed image $\boldsymbol{u}$ and the data $\boldsymbol{g}$ to at least $K = (2\,N_x - 1) \times (2\,N_y - 1) \times (2\,N_z - 1)$, padding $\boldsymbol{g}$ with zeros and setting $w_k = 0$ in the extended area.

Furthermore, except for very low detector noise, this formulation can account for mixed Poisson + Gaussian noise by approximating it as a non-stationary uncorrelated Gaussian noise [10]:

$$w_k \overset{\text{def}}{=} \begin{cases} \left(\gamma \max(g_k, 0) \, + \, \sigma_k^2\right)^{-1} & \text{if } g_k \text{ is measured,} \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

where γ is the quantization factor of the detector and σ_k^2 is the variance of other approximately Gaussian noises (*e.g.*, read-out noise) at voxel k.

2.2.2. TV regularization term

As most observed objects are smooth with few sharp structures (*e.g.*, edges and spikes), I use as regularizing prior an hyperbolic approximation of the classical 3D total variation[11]:

$$\mathcal{J}_{\mathsf{TV}}(\boldsymbol{u}) = \sum_k \sqrt{\|\nabla_k \boldsymbol{u}\|_2^2 + \epsilon^2} , \tag{4}$$

with ∇_k the finite difference operator to approximate spatial gradient at voxel k. Parameter $\epsilon > 0$ ensures differentiability of $\mathcal{J}_{\mathsf{TV}}$ at 0.

2.3. Step 2: Denoising

The solution $\boldsymbol{u}^+$ of the TV deconvolution step is then "denoised" using learned sparse coding, cleaning artifacts and enhancing resolution especially along the depth axis. Both dictionary learning and sparse coding are done using the SPAM toolbox[1] [6, 7].

2.3.1. Dictionary learning

The dictionary $\mathbf{D}^{[\mathrm{s}=1]}$ is learned on the set $\mathcal{P}^{[1]}$ of all overlapping patches extracted from lateral sections of the 3D deconvolved image $\boldsymbol{u}^+$:

$$\mathbf{D}^{[\mathrm{s}=1]} = \underset{\mathbf{D}\in\mathcal{D}}{\arg\min} \, \min_{\boldsymbol{\alpha}} \left\{ \lambda_0 \sum_\ell |\alpha_{\ell,k}| + \sum_{k,m} \left[q_{k,m} - \sum_\ell D_{\ell,m}^{[\mathrm{s}=1]} \, \alpha_{k,\ell} \right]^2 \right\} , \tag{5}$$

[1] an open source toolbox: http://spams-devel.gforge.inria.fr/index.html

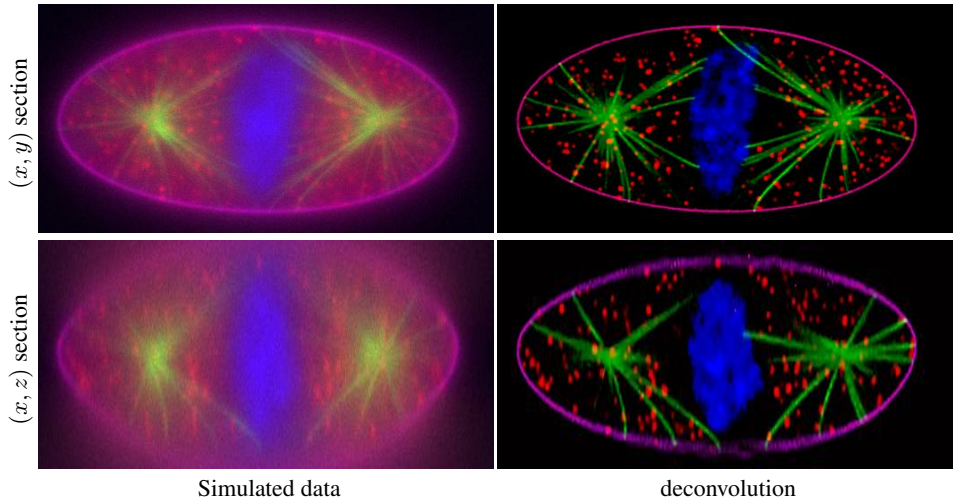


Fig. 2. (x, y) and (x, z) sections of the first *"3D Deconvolution Microscopy Challenge"* data set compared with its deconvolution using the proposed method.

where $\boldsymbol{q}$ is a vector of ℓ_2 normalized patches:

$$q_{k,m} = \frac{p^{[\mathrm{s}=1]}_{k,m}}{\sqrt{\sum_{m'} \left(p^{[\mathrm{s}=1]}_{k,m'}\right)^2}}$$

.

2.3.2. Sparse coding denoising

The learned dictionary is then used to denoise every 2D sections using sparse coding. If the sampling rate differs from one dimension to another, the dictionary is resampled in order to preserve atom physical size along each dimension using interpolation functions $f_{1\to 2}$ and $f_{1\to 3}$: $\mathbf{D}^{[\mathrm{s}=2]} = f_{1\to 2}(\mathbf{D}^{[\mathrm{s}=1]})$ and $\mathbf{D}^{[\mathrm{s}=3]} = f_{1\to 3}(\mathbf{D}^{[\mathrm{s}=1]})$. The denoised image is

$$\boldsymbol{u}^{++}_k = \sum_{k',m,s} R^{[\mathrm{s}]}_{k,k',m} \sum_l D^{[\mathrm{s}]}_{l,m}\, \alpha^{[\mathrm{s}]}_{k',l} \tag{6}$$

$$\boldsymbol{\alpha}^{[\mathrm{s}]} = \arg\min_{\boldsymbol{\alpha}} \sum_m \left(p^{[\mathrm{s}]}_{k,m} - \sum_\ell D^{[\mathrm{s}]}_{\ell,m}\, \alpha^{[\mathrm{s}]}_{\ell,k}\right)^2 + \lambda_s \sum_\ell \left|\alpha^{[\mathrm{s}]}_{\ell,k}\right| . \tag{7}$$

$\mathbf{R}$ is a reconstruction operator which simply copies a weighted version of every synthetic patch k' at its proper position in the 3D image. For instance, if only the central pixels m_0 of the reconstructed patches are averaged:

$$R_{k,k',m,s} = \begin{cases} 1/3 & \text{if } k = k' \text{ and } m = m_0, \\ 0 & \text{otherwise,} \end{cases} , \tag{8}$$

as each pixels are at the center of only 1 patch in each of the 3 subspaces.

2.4. Step 3: Deconvolution using quadratic prior

As expected, the solution of the step 2 $\boldsymbol{u}^{++}$ may be negative and yields to a model $(\mathbf{H}\,\boldsymbol{u}^{++})$ which is too far from the actual data as attested by the value of $\mathcal{J}_{\mathsf{data}}(\boldsymbol{u}^{++})$. As a consequence, a final deconvolution step using $\boldsymbol{u}^{++}$ as a quadratic prior was added:

$$\boldsymbol{u}^{+++} = \arg\min_{\boldsymbol{u}\geq 0} \left(\mathcal{J}_{\mathsf{data}}(\boldsymbol{u}) + \mu_2 \left\|\boldsymbol{u} - \boldsymbol{u}^{++}\right\|_2^2\right) , \tag{9}$$

3. CHOICE OF THE PARAMETERS

For both deconvolution steps (step 1 and 3), few parameters were needed, namely the Gaussian noise variance in each channel, the hyper-parameters μ_1 and μ_2 and the parameter ϵ of the TV hyperbolic approximation.

For the learned sparse coding step, the main parameters are set according to Mairal prescriptions [6, 7]:

- size of the patch $M^{[1]} = 9 \times 9$,
- number of elements of the dictionary $L = 500$,
- hyper-parameter $\lambda_s = 1/\sqrt{M^{[\mathrm{s}]}}$.

In addition to these parameters, there is however a lot of possible variations of the presented method. For instance any other deconvolution method can be used as step 1, orthogonal matching pursuit can be preferred to the LASSO to solve Eq. (6) and/or Eq. (5). Such questions should be assessed in further more extensive studies.

4. RESULTS

4.1. Simulations

This method was first tested on synthetic micrographs simulated for "the 3D Deconvolution Microscopy Challenge" held during ISBI 2013. This dataset is composed of four channels with $1024 \times 512 \times 256$ voxels. The four channels of this dataset were deconvolved independently using an aberration free diffraction PSF. The deconvolution results, presented in Fig. 2, show the strong improvement both in term of resolution and denoising. The increase of resolution is particularly spectacular along the depth axis for sharp structures like the filaments in green. With these results, the presented method reached the first rank at the "the 3D Deconvolution Microscopy Challenge". Unfortunately, ground truth and the quantitative figures of merit of each participant of this challenge are still undisclosed.

4.2. Real data

I have demonstrated the robustness of the presented method by processing experimental data shared by Cazares *et al.* on the *cell image library* website[2]. This dataset consists on an observation of microtubules in a Drosophila S2 cell labeled Alexa Fluor 488 and observed with a Zeiss Elyra structured illumination microscope (SIM) with a $\mathrm{NA} = 1.4$ objective. It is composed of a 3D widefield image and a super-resolution SIM image that can be used as a ground truth. Both micrographs have $(1024 \times 1024 \times 44)$ voxels of size $40 \times 40 \times 110\,\mathrm{nm}$. The PSF, parametrized by its aberrations, was estimated by deconvolving the widefield data by the super-resolution SIM data using the PSF estimation stage of the blind deconvolution algorithm presented in [8].

The data, the deconvolution result and the structured illumination image are shown on Fig. 3. Compared to widefield, the deconvolution result confirms the effectiveness of my method to increase the resolution especially along the depth dimension. Furthermore, even if much more details can be seen on SIM observation that still have a better resolution, the deconvolution result seems less noisy.

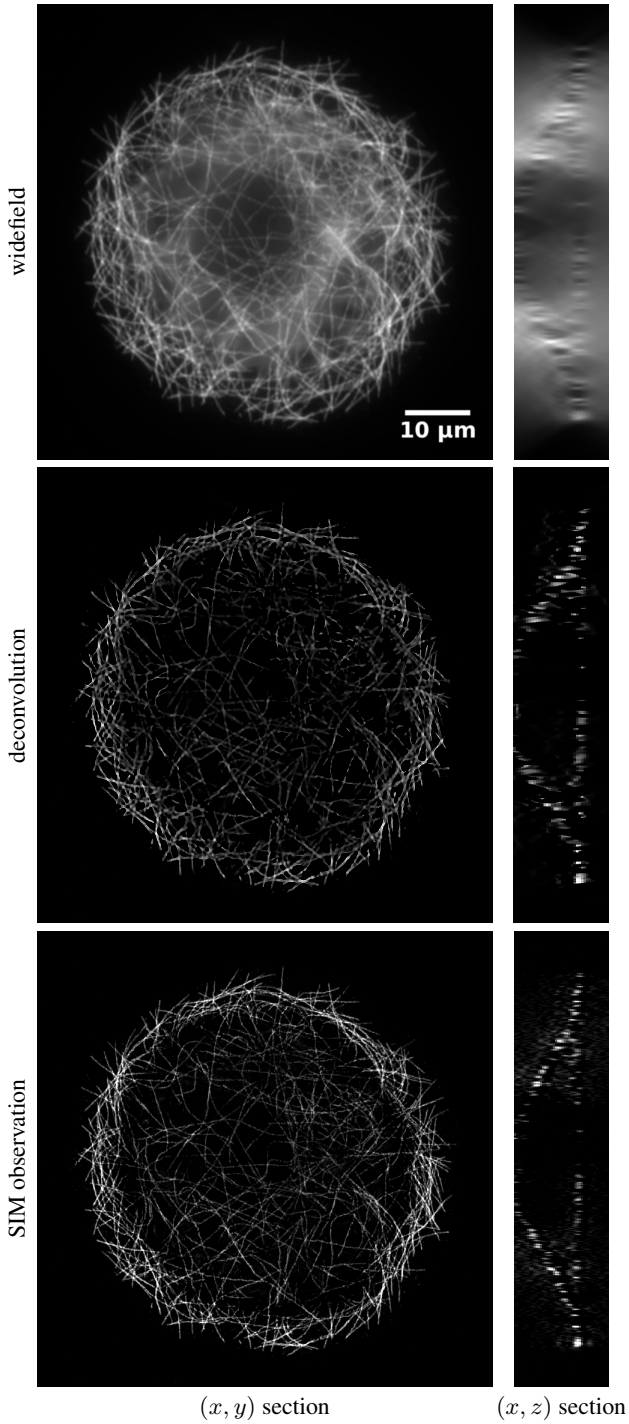


Fig. 3. Widefield observation and its deconvolution compared to the super-resolution image given by structured illumination (SIM) on Drosophila S2 cell.

5. REFERENCES

[1] D.A. Agard, "Optical sectioning microscopy: cellular architecture in three dimensions," *Annu. Rev. Biophys. Bio.*, vol. 13, no. 1, pp. 191–219, 1984.

[2] J.B. Sibarita, "Deconvolution microscopy," *Microscopy Techniques*, pp. 1288–1291, 2005.

[3] P. Sarder and A. Nehorai, "Deconvolution methods for 3-d fluorescence microscopy images," *IEEE Sig. Proc. Mag.*, vol. 23, no. 3, pp. 32–45, 2006.

[4] A. Griffa, N. Garin, and D. Sage, "Comparison of deconvolution software in 3d microscopy. a user point of view part 1.," *G.I.T. Imaging & Microscopy*, vol. 1, pp. 43–45, 2010.

[5] A. Griffa, N. Garin, and D. Sage, "Comparison of deconvolution software in 3d microscopy. a user point of view part 2.," *G.I.T. Imaging & Microscopy*, vol. 1, pp. 41–43, 2010.

[6] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," in *Proc. of 26th ICML*. ACM, 2009, pp. 689–696.

[7] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *J. Mach. Learn. Res.*, vol. 11, pp. 19–60, 2010.

[8] F. Soulez, L. Denis, E. Thiébaut, and Y. Tourneur, "Blind deconvolution of 3d data in wide field fluorescence microscopy," in *proc. 9th Int. Symp. on Biol. Imag.*, Barcelona, 2012.

[9] F. Soulez, E. Thiébaut, A. Gressard, R. Dauphin, and S. Bongard, "Heterogeneous Multidimensional Data Deblurring," in *Proc. of the 16th Eur. Sig. Proc. Conf*, Lausanne, 2008.

[10] L. Mugnier, T. Fusco, and J-M Conan, "Mistral: a myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images," *J. Opt. Soc. Am. A*, vol. 21, no. 10, pp. 1841–1854, 2004.

[11] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.

[2] http://www.cellimagelibrary.org/images/36797