



**HAL**  
open science

## Patch-Based Nonlocal Functional for Denoising Fluorescence Microscopy Image Sequences

Jérôme Boulanger, Charles Kervrann, Patrick Bouthemy, Peter Elbau,  
Jean-Baptiste Sibarita, Jean Salamero

► **To cite this version:**

Jérôme Boulanger, Charles Kervrann, Patrick Bouthemy, Peter Elbau, Jean-Baptiste Sibarita, et al.. Patch-Based Nonlocal Functional for Denoising Fluorescence Microscopy Image Sequences. IEEE Transactions on Medical Imaging, 2010, 29 (2), pp.442-453. 10.1109/TMI.2009.2033991 . inria-00541082

**HAL Id: inria-00541082**

**<https://inria.hal.science/inria-00541082v1>**

Submitted on 29 Nov 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Patch-based non-local functional for denoising fluorescence microscopy image sequences

Jérôme Boulanger, Charles Kervrann, Patrick Bouthemy,  
Peter Elbau, Jean-Baptiste Sibarita, Jean Salamero

**Abstract**—We present a non-parametric regression method for denoising 3D image sequences acquired via fluorescence microscopy. The proposed method exploits the redundancy of the 3D+time information to improve the signal-to-noise ratio of images corrupted by Poisson-Gaussian noise. A variance stabilization transform is first applied to the image-data to remove the dependence between the mean and variance of intensity values. This pre-processing requires the knowledge of parameters related to the acquisition system, also estimated in our approach. In a second step, we propose an original statistical patch-based framework for noise reduction and preservation of space-time discontinuities. In our study, discontinuities are related to small moving spots with high velocity observed in fluorescence video-microscopy. The idea is to minimize an objective non-local energy functional involving spatio-temporal image patches. The minimizer has a simple form and is defined as the weighted average of input data taken in spatially-varying neighborhoods. The size of each neighborhood is optimized to improve the performance of the pointwise estimator. The performance of the algorithm (which requires no motion estimation) is then evaluated on both synthetic and real image sequences using qualitative and quantitative criteria.

**Index Terms**—Video-microscopy, fluorescence, image sequence denoising, patch-based approach, Poisson noise, variance stabilization, adaptive estimation, energy minimization.

## I. INTRODUCTION

Fluorescence video-microscopy is an investigation tool used in biology for dynamics analysis at sub-cellular levels. Combined with fluorescent tags such as genetically engineered fluorescent chimeric proteins (e.g. Green Fluorescence Protein GFP), both confocal microscopy and wide-field microscopy allow 3D live protein imaging. Mainly used to analyze isolated cells, confocal microscopy can also be used *in vivo* if combined with endomicroscopy. Unfortunately, when cell viability needs to be preserved and photo-bleaching avoided, light exposure time must be limited, resulting in low signal-to-noise ratios.

While improving the signal-to-noise ratio, denoising may allow us to reduce exposure time and therefore to

open new opportunities in live cell imaging. Moreover, frame rates can be increased without increasing radiation dose, which could be relevant to capture fast events at sub-cellular levels. Finally, if the point spread function of the objective is not affected by denoising, images may still be compatible with a deconvolution process. This allows to significantly increase the performances of deconvolution algorithms for images with low signal-to-noise ratios and as a consequence, the ability to detect and track objects of interest.

Currently, denoising is a widely studied but still open problem in image processing. Many methods have been described in the literature, and a recent comprehensive review can be found in [1], [2]. Methods based on the full knowledge of noise statistics are probably the most efficient. In fluorescence video-microscopy, it is established that the low level of fluorescence is related to a limited number of photons that can be modeled as a Poisson process. Besides, additive electronic noise is usually present even if a cooling system is used on the detector. Therefore the resulting images are assumed to be contaminated by a combination of Poisson and Gaussian noise. Several approaches have been introduced to deal with such signal-dependent noise. In [3], the authors proposed a maximum penalized likelihood estimator for Poisson noise removal in very low count situations. The problem is more challenging for Poisson-Gaussian noise and another line of work consists in stabilizing the noise variance using ad-hoc transforms. The more common transform is the so-called Anscombe transform [4] designed for Poisson noise. This transform was further generalized to Poisson-Gaussian noise [5], with satisfying results if the number of counts is large enough and more recently for "clipped" (under- and over-exposure) raw-data [6]. In the case of very low count situations ( $\leq 1$  photons in average), the more sophisticated Fisz transform allows one to better stabilize Poisson noise [7], [8]. Finally, local estimation of image-dependent noise statistics (assumed to be locally Gaussian) has also been investigated, especially in the case of adaptive Wiener filtering [9]–[11].

Denoising temporal sequences is even more com-

plex since there are currently no satisfactory methods for processing fluorescence video-microscopy 3D image sequences contaminated by Poisson-Gaussian noise. Most of them only restore each frame separately without using the temporal redundancy of image series [12], [13]. When temporal coherence is exploited, it is usually recommended to consider a motion estimation/compensation stage as proposed for video denoising [14]–[17] and, for instance, for low-dose fluoroscopy image sequence filtering [11]. This is especially true for real-time imaging applications. Thus, Kuznetsov *et al.* recently proposed to use a temporal Kalman-Bucy filter to improve the quality of video-microscopy images [18]. The main difficulty in video-microscopy is to estimate the motion of small and similar objects moving with high velocity in the image sequence. To overcome this problem, sophisticated methods (see [1]) but designed for still images have been adapted to videos. Wavelet shrinkage [19], [20], Wiener filtering [21] or PDE-based methods [22] are typical examples of such methods. Some of them have been successfully adapted to video-microscopy [23], [24]. Recently, an extension of the non-local means filter [1] also related to the universal denoising (DUDE) algorithm [25] and the entropy-based UINTA filter [26], has been proposed to process image sequences. It assumes that an image sequence contains repeated patterns [27]. Noise can then be reduced by averaging data associated to the more similar patches in the image sequence. Finally, patch-based approaches are now very popular in texture synthesis [28], inpainting [29] and video completion [30].

Nevertheless, searching similar examples in the whole image for denoising with the non-local means filter, is untractable in practice in 2D, and unrealistic for video sequences. As a consequence, a variant of this filter has been recently proposed in [31] in which the authors use a pre-classification of the pixels of the sequence in order to speed up the denoising procedure. Another improvement introduced in [32] consists in collecting similar patches to build 3D arrays. A unitary transform and a hard-thresholding are then applied to remove noise. In the meanwhile, a general modeling framework based on signal theory and machine learning has been proposed by Elad *et al.* for image and video sequence analysis. The authors assume that the image is sparsely represented over an over-complete dictionary of atoms that are either fixed (e.g. DCT) or learned from exemplar patches [33], [34]. The approximation problem is then equivalent to the minimization (using a K-SVD algorithm) of an energy functional involving a data term and a penalty term that encodes sparsity [35]. This method is able to produce impressive image denoising results, including

on image sequences, but requires intensive minimization procedures and the adjustment of several parameters.

Unlike the previous patch-based approaches [27], [31], [34], [36], we present in this paper a space-time patch-based adaptive statistical method for 3D+time video-microscopy image sequence restoration. As already mentioned, patch-based methods have been proposed for denoising image sequences, but, to our knowledge, only anisotropic diffusion and wavelet shrinkage have been applied to 2D+time fluorescence video-microscopy [23], [24]. The main features of the proposed method have already been presented in a discrete setting at the IEEE-ISBI'08 conference [37]. In our approach, we propose first a variance stabilization step to be applied to the data in order to obtain independence between the mean and the variance. Second, we consider spatio-temporal neighborhoods to restore series of 3D images as already proposed for 2D image sequences in [36]. Our method is based on the minimization of an energy functional while exploiting image patches. The minimizer of this energy functional established in a continuous setting has a simple form and corresponds to a weighted average of intensity values taken in spatially (and temporally) varying neighborhoods. The neighborhood size is adapted on-line to improve the performance (in the sense of the  $L_2$  risk) of the pointwise estimator. No learning step or wavelet decomposition is required. Also, no motion estimation is involved as originally described in [36]. Finally, the designed algorithm comprises only a few parameters which are easily calibrated.

The remainder of this paper is organized as follows. In Section II, we introduce the denoising problem in fluorescence video-microscopy. In Section III, we first present the generalized Anscombe transform and detail an original approach to estimate its parameters and correct the induced bias. Then, we introduce the space-time patch-based estimator. In Section IV, we demonstrate the performance of the algorithm (controlled by a small number of parameters) on both synthetic and real video-microscopy images and image sequences.

## II. PROBLEM STATEMENT

In this section, we present a general framework for image sequence analysis in wide-field or confocal microscopy. Our study is limited to the restoration of artifacts due to random noise. We do not consider the issue of correcting the signal distortions due to diffraction (e.g. deconvolution problem) but we will later show the compatibility of the proposed method with a deconvolution post-processing step.

Acquired images correspond to stacks of 10 to 60 slices with an axial resolution (depth) lower than the

lateral one. Anisotropy in 3D microscopy can be an issue for 3D wavelet methods, especially for processing stacks with a limited number of slices due to boundary effects. The processed images depict tagged proteins appearing as bright particles of size 3 to 10 pixels and moving with speeds ranging from 1 to 10 pixels per frame. The small amount of light collected by sensors and thermal agitation in electronic components induce a Poisson-Gaussian noise. Accordingly, we assume the following affine stochastic model:

$$Z_i = g_0 N_i + \varepsilon_i, \quad (1)$$

where  $Z_i \triangleq Z(x_i)$  is the observation at space-time location  $x_i \in \mathbb{R}^d$ ,  $i \in \{1, \dots, n\}$  and  $d$  the dimension of the space-time domain. The gain of the overall electronic system is denoted  $g_0$ . The number  $N_i$  of collected photoelectrons at pixel  $x_i$  is a random variable assumed to follow a Poisson distribution of parameter  $\theta_i \triangleq \theta(x_i)$  with density:  $p(N_i) = \frac{\theta_i^{N_i} e^{-\theta_i}}{N_i!}$ . Finally, the dark current is treated as a Gaussian white noise of mean  $\mathbb{E}[\varepsilon_i] = m$  and variance  $\text{Var}[\varepsilon_i] = \sigma_\varepsilon^2$ . In our model, the two random variables  $N_i$  and  $\varepsilon_i$  are independent. Finally, we denote  $f_i \triangleq f(x_i) = g_0 \theta(x_i) + m$ .

In this paper, we consider the problem of estimating  $f_i$  at each pixel  $x_i$  from noisy data  $Z_i$ . A root-unroot strategy [38] is considered to deal with the Poisson-Gaussian noise context while a patch-based functional yields an estimator of the intensity value whose parameters are estimated in an iterative fashion.

### III. PROPOSED METHOD

#### A. Noise variance stabilization

1) *Definition*: The Anscombe transform is the more commonly-used transform for stabilizing the variance of Poisson noise [4]. Murtargh *et al.* considered a more general Anscombe transform (GAT) for Poisson and Gaussian noise [39]. Using the notation introduced in (1), the GAT can be expressed as:

$$\mathcal{T}_{GA}(Z_i) = \frac{2}{g_0} \sqrt{g_0 Z_i + \frac{3}{8} g_0^2 + \sigma_\varepsilon^2 - g_0 m}. \quad (2)$$

Note that variance stabilization and skewness correction are incompatible.

2) *Parameter estimation*: In contrast to the usual parameter-free Anscombe transform, the GAT requires the setting (or the estimation) of a small set of parameters,  $g_0$ ,  $\sigma_\varepsilon^2$  and  $m$ , related to the acquisition system. In [40], the authors proposed a bias-variance trade-off criterion to determine the parameters of their multi-scale variance stabilization transform. However, they do not provide the method to estimate the parameters  $g_0$ ,  $\sigma_\varepsilon^2$

and  $m$ . Nevertheless, Starck *et al.* proposed in [41] an iterative algorithm to estimate the gain  $g_0$  and the dark current parameters from images.

Instead, we propose an approach based on a linear regression in the 2D-space ( $\mathbb{E}[Z_i]$ ,  $\text{Var}[Z_i]$ ). This method has been previously sketched in [42] and we provide here additional details and some improvements. A similar approach has been since described in [24], [43]. From (1), we have

$$\begin{cases} \mathbb{E}[Z_i] &= g_0 \theta_i + m, \\ \text{Var}[Z_i] &= g_0^2 \theta_i + \sigma_\varepsilon^2. \end{cases} \quad (3)$$

which yields

$$\text{Var}[Z_i] = g_0 \mathbb{E}[Z_i] + \sigma_\varepsilon^2 - g_0 m. \quad (4)$$

It follows that a linear regression in the 2D-space ( $\mathbb{E}[Z_i]$ ,  $\text{Var}[Z_i]$ ) provides an estimation of the two parameters  $g_0$  and  $e_{DC} = \sigma_\varepsilon^2 - g_0 m$ . Accordingly, (2) can be written as

$$\mathcal{T}_{GA}(Z_i) = \frac{2}{g_0} \sqrt{g_0 Z_i + \frac{3}{8} g_0^2 + e_{DC}}. \quad (5)$$

In order to get uncorrelated estimates of the local mean and of the local variance, it is crucial to partition the space-time volume into non-overlapping blocks. Instead of defining in advance the size of these blocks, we propose to divide the image using a quadtree/octree segmentation procedure. Each region is recursively divided into four/eight smaller regions if the variance of the data  $Z_i$  in the current region is not explained by the variance of the noise. The variance in a region  $R$  containing  $|R|$  pixels is given by:  $S_Z(R) = \sum_{i \in R} (Z_i - \frac{1}{|R|} \sum_{j \in R} Z_j)^2 / (|R| - 1)$ . The variance of the noise is defined by  $S_\varepsilon(R) = \sum_{i \in R} (r_i - \frac{1}{|R|} \sum_{j \in R} r_j)^2 / (|R| - 1)$  where the pseudo-residuals  $r_i$  are defined by (see [44]):

$$r_i = \frac{1}{\sqrt{l^2 + l}} \Delta Z_i. \quad (6)$$

Here  $\Delta Z_i$  denotes the Laplacian operator involving  $l = 2d + 1$  surrounding pixels and is defined for a  $d$ -dimensional space as:  $\Delta Z_i = l Z_i - \sum_{j=1}^d (Z(x_j + s_j) + Z(x_j - s_j))$  with  $s_j$  a vector whose  $j$ th coordinate is 1 and the other 0. Furthermore, a Fisher test is used to compare the two variances:  $\min(S_Z(R), S_\varepsilon(R)) / \max(S_Z(R), S_\varepsilon(R)) \leq T_{\alpha_F, |R|-1}$ . The threshold  $T_{\alpha_F, |R|-1}$  corresponds to the  $\alpha_F$ -quantile of F-distribution with  $|R| - 1$  degrees of freedom. This procedure results in a partition of the image into regions with homogeneous variance. Figure 1 shows an example of such an image partition. Finally, estimates of pairs of local mean and variance can be then robustly estimated within these regions. The mean can be estimated using a

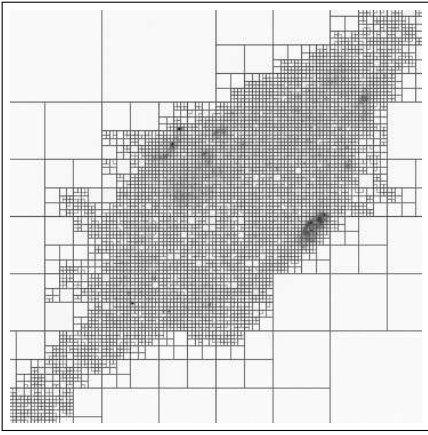


Fig. 1. Partition of the image domain using a quadtree segmentation based on the comparison of the local variance of the image and the local variance of the noise. The image corresponds to the exposure time of 500ms as shown in Fig. 8.

robust M-estimator (using a Leclerc influence function) [45] while an estimate of the variance of the noise is provided by the “Least Trimmed Square” robust estimator [46].

Given empirical estimates of the mean and the variance, a robust linear regression provides the values of parameters  $g_0$  and  $e_{DC}$ . The Generalized Anscombe Transform is then applied to the input data  $\{Z_i\}_{i \in [1, n]}$  to produce new input data  $\{Y_i = \mathcal{T}_{GA}(Z_i)\}_{i \in [1, n]}$  with Gaussian statistics. Finally, in order to be able to generate images with the same noise signature defined by the triplet  $(g_0, \sigma_\varepsilon, m)$ , one has to estimate first the parameters of the dark current  $\sigma_\varepsilon$  and  $m$ . In most images, these two parameters can be deduced from the variance and the intensity values corresponding to the darker regions.

3) *Image quality assessment*: Evaluating the image quality, is an important step in video-microscopy as it will allow to measure the errors involved in the quantification steps [47]. It is worth noticing that the Generalized Anscombe transform provides a way to evaluate the image quality of acquired images. Once stabilized, the noise variance is expected to be homogeneous in the whole image domain and equal to 1. Hence, we can define the following Poisson Peak Signal to Noise Ratio:

$$\text{PPSNR}(Z) = 20 \log_{10} \left( \max_i \{\mathcal{T}_{GA} Z_i\} - \min_i \{\mathcal{T}_{GA} Z_i\} \right)$$

Instead of the image contrast, this measure could also involve the contrast of objects using a background subtraction method in the same fashion than the signal-to-noise S/N ratio introduced in [48] for astronomical images and nowadays used in microscopy [47]. Finally, this approach provides a fully automatic quantification of the image quality.

4) *Un-biased inverse GAT*: After variance stabilization, one can apply an algorithm designed for Gaussian noise to the transformed data  $\{Y_i\}_{i \in [1, n]}$  and get an estimate  $\hat{u}$  of the underlying function  $u$  defined on the image domain  $\Omega \subset \mathbb{R}^d$ , with  $d$  the space-time dimension. At location  $x_i \in \Omega$  we have  $Y_i = u(x_i) + \xi_i$  with  $\xi_i$  a Gaussian centered white noise of variance 1. Then inverting the Generalized Anscombe Transform yields to an estimate  $\hat{f} = TGA^{-1}(\hat{u})$  of the function  $f$ . However, this procedure would introduce an additive bias. When the number of counts is high and when the number of samples is large enough, the bias tends to  $1/4$ . Figure 2 illustrates this effect on the estimation of the count of a Poisson distributed random variable  $\theta$ . The bias exhibits a behavior that can be heuristically approximated by  $(1 - \exp(-1.3\theta))/4$  where the coefficient 1.3 has been estimated from the simulation shown in Fig. 2. Since  $\theta$  is unknown, an iterative procedure is used to estimate the bias correction operator  $C_{AT}(f)$  defined as:

$$C_{AT}(f)(x_i) = \frac{1}{4} \left( 1 - e^{-1.3\hat{f}(x_i) + C_{AT}(f)(x_i)} \right)$$

where  $\hat{f}(x_i)$  is the value obtained by directly inverting the Anscombe transform at point  $x_i$ . Finally, the unbiased estimate is given by  $\hat{f}_{unbiased} = \hat{f} + C_{AT}$ . This experiment contradicts the conclusion drawn in [5]. In particular, the Anscombe transform performs well for  $\theta > 3$  instead of  $\theta > 30$ , which is reasonable for our application.

### B. Patch-based space-time estimation

In this section we first extend the continuous non-local patch-based functional introduced by Kindermann et al. in [49]. Given its fixed point solution we derive another functional also related to [50]. We finally present a discretization of the minimizer and a method to estimate its parameters.

1) *Non-local functional of Kindermann, Osher and Jones*: Kindermann et al. introduced in [49] a non-local patch-based functional for denoising and deblurring images. This functional is built upon a new norm which measures the degree of similarities between patches. We propose to extend it as follows:

$$J(u, u_0) = \frac{\lambda}{\sigma^2} \int_{\Omega} (u(x) - u_0(x))^2 dx + \int_{\Omega^2} \phi \left( \int_{\Omega} G(t) \frac{(u(x+t) - u(y+t))^2}{Q(x+t, y+t)} dt \right) K \left( \frac{\|x-y\|}{h(x, y)} \right) dy dx, \quad (7)$$

where  $u$  is the function to estimate defined on the image domain  $\Omega \subset \mathbb{R}^d$  and  $d$  the dimension of the space-time domain. The function  $u_0$  represents the initial data  $Y$  (i.e.  $u_0 = u + \xi$  where  $\xi$  is the noise as defined in

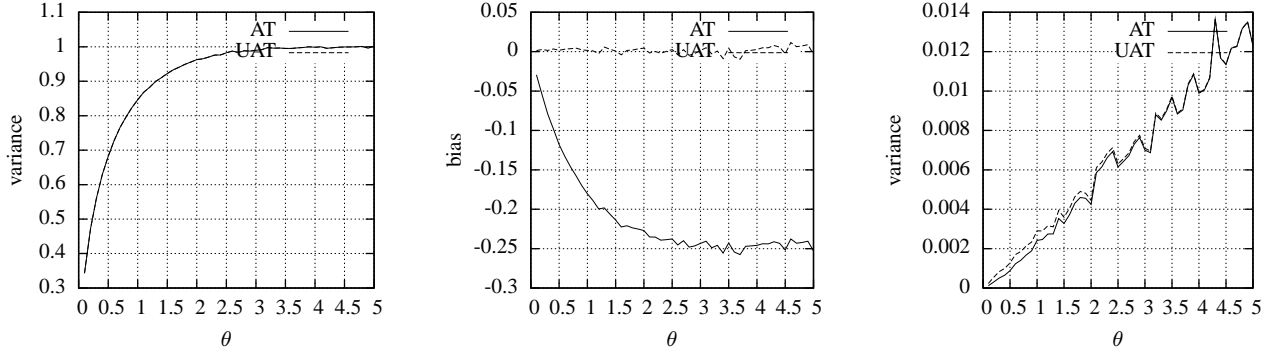


Fig. 2. Analysis of the bias of the Anscombe transform for the mean estimator of 400 Poisson distributed samples in the range  $[0, 5]$ . On the left, the stabilized variance is displayed for the Anscombe Transform (AT) and the unbiased Anscombe transform (UAT) which are by definition the same. The middle graph shows the bias of the two estimators computed from 400 trials. On the right, the variances of the estimators are displayed.

Section III-A4). The function  $\phi$  is a  $\mathbb{R} \rightarrow \mathbb{R}$  differentiable function (typically  $\phi(x) = 1 - e^{-x}$ ). The proposed extension lays in the introduction of the locally variable bandwidths defined by the two functions  $\Omega \times \Omega \mapsto \mathbb{R}$ :  $Q$  and  $h$ . To be able to derive a fixed point iteration, one can show that the symmetry of these bandwidths i.e.  $Q(x, y) = Q(y, x)$  and  $h(x, y) = h(y, x)$ , is needed.

The fixed point equation for minimizing (7) has the following form:

$$\hat{u}(x) = \frac{u_0(x) + \frac{2\sigma^2}{\lambda} \int_{\Omega^2} A(x, y, z) u(y) dz dy}{1 + \frac{2\sigma^2}{\lambda} \int_{\Omega^2} A(x, y, z) dz dy} \quad (8)$$

where

$$A(x, y, z) = \frac{G(z)}{Q(x, y)} K\left(\frac{\|x-y\|}{h(x-z, y-z)}\right) \times \phi' \left( \int_{\Omega} G(t) \frac{(u(x+t-z) - u(y+t-z))^2}{Q(x+t-z, y+t-z)} dt \right). \quad (9)$$

The convergence of the fixed point iteration is not guaranteed. We can also point out the fact that the minimizer of the functional (7) involves overlapping patches which is an original feature compared to other patch-based variational approaches [50]–[53].

2) *Proposed functional*: Instead of (7), we propose to minimize the following functional (see also [50]):

$$J(u, u_0) = \int_{\Omega} \left( \frac{\int_{\Omega} B(x, y, z) u_0(y) dz dy}{\int_{\Omega} B(x, y, z) dz dy} - u(x) \right)^2 dx \quad (10)$$

where

$$B(x, y, z) = \frac{G(z)}{Q(x, y)} K\left(\frac{\|x-y\|}{h(x-z, y-z)}\right) \times \phi' \left( \mathbb{E} \left[ \int_{\Omega} G(t) \frac{(u_0(x+t-z) - u_0(y+t-z))^2}{Q(x+t-z, y+t-z)} dt \right] \right). \quad (11)$$

This expression relies on the expectation of the distance between patches instead of the distance itself making it less sensitive to noise.

The minimizer of the functional (10) is trivial since the function  $u$  does not appear in the first term. On the other hand, the calculation of the expectation in (11) makes the evaluation of  $B(x, y, z)$  difficult. However, if  $\mathbb{E}[u_0(x)] = u(x)$  and  $\text{Var}[u_0(x)] = \sigma^2$ , we can use the following identity (see [1]):

$$\mathbb{E} \left[ \int_{\Omega} G(t) \frac{(u_0(x+t) - u_0(y+t))^2}{Q(x+t, y+t)} dt \right] = \int_{\Omega} G(t) \frac{(u(x+t) - u(y+t))^2 + 2\sigma^2}{Q(x+t, y+t)} dt \quad (12)$$

Finally, since  $J$  is positive, we have:

$$\begin{aligned} \hat{u}(x) &= \arg \min_u J(u, u_0) \\ &= \arg \min_u \frac{\int_{\Omega^2} \tilde{B}(x, y, z) u_0(y) dz dy}{\int_{\Omega^2} \tilde{B}(x, y, z) dz dy}, \end{aligned} \quad (13)$$

with

$$\begin{aligned} \tilde{B}(x, y, z) &= \frac{G(z)}{Q(x, y)} K\left(\frac{\|x-y\|}{h(x-z, y-z)}\right) \\ &\times \phi' \left( \int_{\Omega} G(t) \frac{(u(x+t-z) - u(y+t-z))^2 + 2\sigma^2}{Q(x+t-z, y+t-z)} dt \right). \end{aligned} \quad (14)$$

3) *Numerical aspects and discretization*: For the implementation of the estimator defined by equation (13), we consider a discrete setting. We can also initialize the fixed point iteration using the data obtained after variance stabilization and set  $u_0 = Y$ . We have thus the following expression for the estimator:

$$\hat{u}_i = \sum_{j=1}^n \sum_{k=1}^n \omega_{ijk} Y_j \quad (15)$$

where  $\omega_{ijk} = \tilde{B}(x_i, x_j, x_k) / (\sum_{j=1}^n \sum_{k=1}^n \tilde{B}(x_i, x_j, x_k))$  and  $\hat{u}_i$  denotes the fixed point solution at pixel  $x_i$ . We can also compute the following approximation for the variance of this estimator:

$$\hat{v}_i = \sigma^2 \sum_{j=1}^n \sum_{k=1}^n \omega_{ijk}^2. \quad (16)$$

These two expressions are almost equivalent to the estimator (and its variance) introduced in [37]. As a trade-off between computational efficiency and simplicity, the kernel  $G$  is defined as the indicator function on the interval  $[-p/2, p/2]^d$ . In addition, we define the ‘‘tonal’’ bandwidth as  $Q(x_i, x_j) = (\hat{v}_i + \hat{v}_j)/(\lambda_\alpha \hat{v}_i \hat{v}_j)$ . Given the shape of  $G$ , the parameter  $\lambda_\alpha$  is related to an  $\alpha$ -quantile of the  $\chi^2$  distribution whose number of degrees of freedom is given by  $n_G - 1$  where  $n_G$  is the number of points lying in the support of the kernel  $G$ . This definition of  $G$  fulfills the condition of symmetry. Under some assumption, the fixed-point iterations converge relatively fast and few iterations are used in practice.

4) *Space-time bandwidth selection:* We define now the bandwidth  $h(x, y)$  of the kernel  $K$  as  $h(x, y) = \min(h(x), h(y))$  and consider the estimation of  $h_i \triangleq h(x_i)$  for each point  $x_i$  of the image sequence. We would like to select the bandwidth minimizing the mean square risk of the proposed estimator defined as  $\mathcal{R}(\hat{u}_i, u(x_i)) = \mathbb{E}[(\hat{u}_i - u(x_i))^2]$ . This risk can be decomposed as the sum of the squared bias and the variance. The bias can not be directly estimated because it depends on the unknown function  $u$ . However we can use an upper bound for the squared bias  $b_i^2$  term and derive the following property for the optimal estimator  $\hat{u}_i^*$  [54]:

$$\frac{(b_i^*)^2}{v_i^*} = \frac{d^2}{4} \triangleq \gamma^2, \quad (17)$$

where  $d$  is the dimension of the space-time domain. Expression (17) does not depend on image regularity. Following the Lepskii’s principle [55], we exploit this property to minimize the  $L_2$  risk  $\mathcal{R}(\hat{u}_i, u(x_i))$ . The idea is to design a sequence of increasing bandwidths  $\mathcal{H}_i = \{h_i^\ell, \ell \in \{0, \dots, L-1\} : h_i^{\ell-1} \leq h_i^\ell\}$ . Assuming that the variance  $v_i^\ell$  is a decreasing function of  $h$  the number of samples taken into account is progressively increased to reduce the estimator variance while controlling the estimator bias. Formally, the so-called ‘‘bias-variance trade-off’’ corresponds to the following inequality:

$$h_i^* = \sup_{h_i^\ell \in \mathcal{H}_i} \{|b_i^\ell|^2 \leq \gamma^2 v_i^\ell\}. \quad (18)$$

This stepwise procedure provides a reasonable estimate of the bandwidth minimizing the local quadratic risk within the pre-defined set  $\mathcal{H}$ . Since the bias  $b_i^\ell$  is unknown, we consider instead a weaker ‘‘oracle’’ to detect the optimal bandwidth for smoothing (see [56], [57]):

$$h_i^* = \sup_{h_i^\ell \in \mathcal{H}_i} \{\ell' < \ell : |\hat{u}_i^{\ell'} - \hat{u}_i^\ell|^2 \leq \rho v_i^{\ell'}\} \quad (19)$$

where  $\rho$  is a positive constant (we choose  $\rho = 8$ , see [58]). The design of a sequence of increasing

bandwidths is now required. However, in the case of image sequences, the relationship between the temporal and spatial dimensions is related to the object size and movement, which are both unknown. Accordingly, the space and time bandwidths should be considered independently. For this reason, we decide to increase alternatively the size of the support of  $K$  using two distinct radii. We note respectively  $h_s$  and  $h_t$  the spatial and temporal neighborhoods which can vary from one point to another. It is worth noting that, unlike [57], the sequence of shape of  $K$  is not known in advance since we consider two parameters  $h_s$  and  $h_t$ . In our experiments, we use a dyadic scale in space and a linear scale in time to achieve a compromise between accuracy and computational efficiency.

5) *Wiener filter:* In [59], a Wiener filter is used to combine estimates obtained at each iteration while in [1], the same approach is used to recover details after filtering. We have also observed some improvement using such approach and propose to filter at each iteration the successive estimates:

$$(\hat{u}_i^\ell)^{\text{Wiener}} = \left( \frac{\hat{u}_i^\ell}{\hat{v}_i^\ell} + \frac{\hat{u}_i^{\ell-1}}{\hat{v}_i^{\ell-1}} \right) \left( \frac{\hat{v}_i^\ell \hat{v}_i^{\ell-1}}{\hat{v}_i^\ell + \hat{v}_i^{\ell-1}} \right) \quad (20)$$

and

$$(\hat{v}_i^\ell)^{\text{Wiener}} = \frac{\hat{v}_i^\ell \hat{v}_i^{\ell-1}}{\hat{v}_i^\ell + \hat{v}_i^{\ell-1}}. \quad (21)$$

Actually, isolated and unaltered pixels in the restored image can be slightly modified using this filtering, which enhances image quality.

6) *Patch pre-selection:* Finally, we propose to extend the patch pre-selection related to [31], [60] to reduce the computational load and in the meanwhile improve the results. Thus, the weights  $\omega_{ijk}$  in (15) and (16) are set to 0 if

$$\frac{(G * \hat{u}_i^\ell - G * \hat{u}_j^\ell)(G * \hat{v}_i^\ell G * \hat{v}_j^\ell)}{2(G * \hat{v}_i^\ell + G * \hat{v}_j^\ell)} > \eta_1 \quad (22)$$

and

$$\frac{\max(G * \hat{v}_i^\ell, G * \hat{v}_j^\ell)}{\min(G * \hat{v}_i^\ell, G * \hat{v}_j^\ell)} > \eta_2, \quad (23)$$

where  $*$  denotes the convolution operator,  $G$  remains the same kernel than in (7) and  $\eta_1$  and  $\eta_2$  are respectively two thresholds (with some approximations) related to a quantile of the Normal distribution and to a quantile of the F-distribution.

In the following experiments, Wiener filtering and patch pre-selection were used to speed-up the computation time and enhance the image quality.

## IV. EXPERIMENTS

### A. Synthetic image sequence

In order to test the proposed method, we have generated synthetic image sequences representing moving tagged vesicles. Using this procedure, we aim to analyze the influence of the generalized Anscombe transform on the final result and to demonstrate that the proposed space-time adaptive method is competitive when compared to the state-of-the-art methods.

First, we have created a synthetic image sequence showing moving objects superimposed on a static background. The true image sequence is composed of 50 frames of 16 bits 3D volumes of  $256 \times 256 \times 10$  voxels. The background is generated using two or three Gaussian profiles of radius 20 pixels at random locations. The background is an essential component of the photometric dynamic of images and thus will probably alter the stabilization process. Typically, the background may be associated to auto-fluorescence within the cell as well as the non specific accumulation of fluorescent tags on organelles. The flux of photo-electrons related to this component ranges from 10 to 2000 photo-electrons per pixel. In addition, 256 spots are drawn as 3D Gaussian functions of radius 2 pixels and of intensity 200 photo-electrons. The movements of objects are assumed to be described by a Gaussian random walk of standard deviation of 3 pixels. A Poisson noise is generated from this image of flux. Then a gain  $g_0 = 0.4$  is applied and finally the dark current is simulated with a Gaussian noise of mean  $m = 100$  and a standard deviation  $\sigma_\epsilon = 4$ . All these values have been obtained by statistical analysis of photometric properties observed in real image sequences. The synthetic image sequence is composed of small spots with intensities of 70 gray levels above the background level, and of 4 large blobs with a maximal intensity of about 900. The slice #5 extracted from a volume at time  $t = 25$  of the simulated (noise free) ground truth and the corresponding noisy slice are shown respectively in Fig. 3(a) and (b).

A scatter plot of the estimated mean and noise variance is shown in Fig. 4(a). The regression line for the first image of the sequence is estimated as  $\widehat{\text{Var}}[Z_i] = 0.407 \widehat{\mathbb{E}}[Z_i] - 24.44$ , while the true equation is  $\text{Var}[Z_i] = 0.4 \mathbb{E}[Z_i] - 24.0$ . We can analyze the accuracy of the estimation by considering the next volumes of the sequence. We found that the mean of  $g_0$  is 0.408 and the standard deviation is  $6.79 \cdot 10^{-3}$ . For the parameter  $e_{DC}$ , the mean is  $-24.31$  and the standard deviation is 0.879. Accordingly, we can conclude that, for this simulation, the parameters of the generalized Anscombe transform have been satisfyingly estimated. In addition,

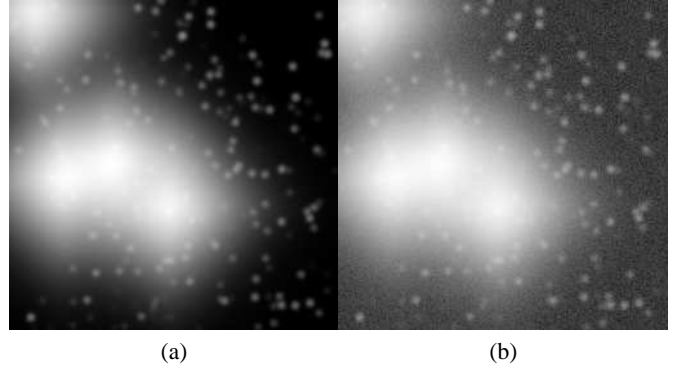


Fig. 3. Volume of  $256 \times 256 \times 10$  voxels extracted from a simulated image sequence (slice #5 and time  $t = 25$ ), (a) ground truth (b) noisy image sequence (logarithmic scale).

Fig. 4(b) shows that the variance of the noise has been well stabilized: the noise variance is now 1.001. The width of the cloud of points is related to the estimation errors of the noise variance. However, the global trend is correctly estimated and the noise variance is reliably stabilized.

This simulation shows that our approach is quite effective at stabilizing the noise variance in the case of a Poisson-Gaussian noise. It is fully automatic and fast. The computation time of an unoptimized C++ implementation is about 250ms for a single  $256 \times 256 \times 10$  3D frame  $256 \times 256 \times 10$  on a 1.8Ghz PC. The parameters are estimated for each 3D frames of the sequence and smoothed in time using a moving average in order to take into account the possible variations of the sensor characteristics.

To demonstrate the performance of both the variance stabilization procedure and the 3D+time denoising procedure, we consider three experiments. In experiments A and B, we assume respectively a Poisson-Gaussian noise model and a Gaussian noise model. In experiment C, we assume a Poisson-Gaussian noise model but each volume of the sequence is denoised independently. In these three experiments, we used  $5 \times 5 \times 5$  patches and the algorithm parameters are unchanged.

In order to compare the different methods and noise models, we measured the  $L_\infty$ ,  $L_1$  and  $L_2$  norms (see Table I) between the original sequence  $f$  and the reconstructed image sequence  $\hat{f}$ . The results are reported in Table II and Fig. 5. Finally, we consider the signal-to-noise ratio  $\text{SNR} = 10 \log_{10}(\text{Var}[f] / \|\hat{f}(x) - f(x)\|^2)$ . From a noisy image with  $\text{SNR} = 24.0\text{dB}$  we obtained the following value of SNR: 33.04dB, 31.06dB and 32.55dB respectively for the denoised image sequences corresponding to experiments A, B and C. All the considered metrics show the interest of tacking into



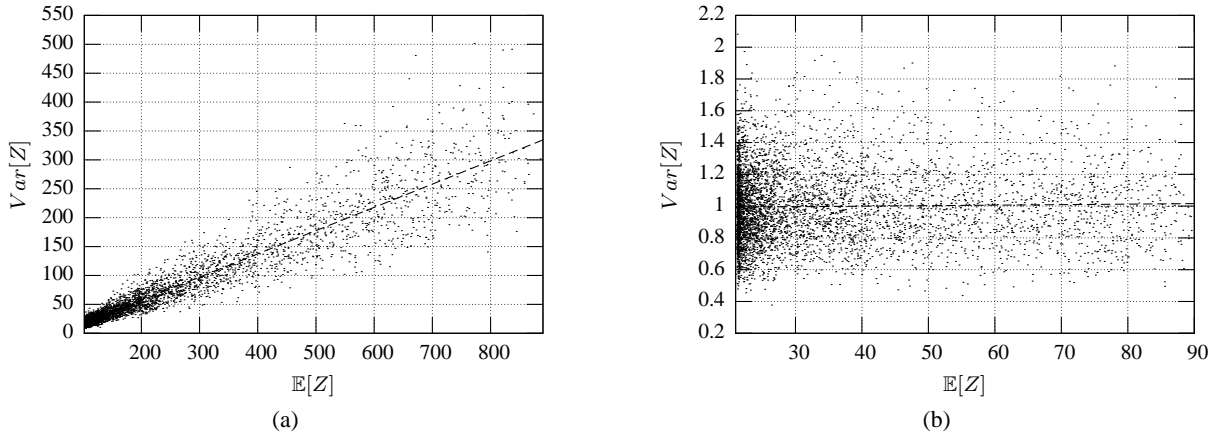


Fig. 4. Noise variance stabilization for a synthetic image sequence. Robust estimation of the local mean  $\mathbb{E}[Z_i]$  and noise variance  $\text{Var}[Z_i]$  (a) before stabilization and (b) after stabilization. Each dot corresponds to a couple  $(\mathbb{E}[Z_i], \text{Var}[Z_i])$  estimated non-overlapping blocks. The dashed line represents the fit of the theoretical model  $\text{Var}[Z_i] = g_0\mathbb{E}[Z_i] + e_{DC}$ . After stabilization, the dependence between the signal intensity and the noise variance is canceled.

$L_p$ norm	
$L_\infty$	$\sup_{x \in \Omega}  f(x) - \hat{v}(x) $
$L_1$	$\int_{x \in \Omega}  f(x) - \hat{v}(x)  dx$
$L_2$	$\int_{x \in \Omega}  f(x) - \hat{v}(x) ^2 dx$

TABLE I  
DEFINITIONS OF  $L_p$  NORMS USED FOR EVALUATION.

account Poisson/Gaussian noise modeling and space-time information.

Moreover the visualization of the sequence restored frame by frame, makes clearly appear a flickering artifact due to the lack of temporal coherence between consecutive images. In Fig. 7 we can notice the differences between experiments A and B. Flickering artifacts are visible in Fig. 7(b) corresponding to experiment B while in Fig. 7(a) the temporal coherence is reinforced. We can also remark that temporal abrupt changes are well preserved. As expected, these experiments visually confirm that considering the whole image sequence provides better results than processing each frame of the sequence independently.

### B. Spatial denoising of real samples using various exposure times

In this section, we consider several spinning disk acquisitions of the same fixed HeLa cell expressing GFP tagged Rab6 proteins. For these experiments, the exposure time varies from 30 to 500ms. The acquired 3D stacks have the size of  $400 \times 400$  voxels. In this case temporal information is not used since the cell is fixed.

Several methods are also applied to these data for comparison:  $3 \times 3$  and  $3 \times 3 \times 3$  median filters combined with the proposed GAT, the multi-scale variance stabilization (MS-VST) approach using a  $7/9$  orthogonal filter [61] and the parameters  $g_0$ ,  $m$  and  $\sigma_0$  estimated as described in Section III-A, the BM3D method [59] combined with the proposed GAT. Finally, we evaluate our method in 2D and 3D.

Results are shown in Fig. 8. The results corresponding to the two median filters, performing badly, are not displayed in order to better focus on the other methods. In this experiment the BM3D method outperforms the other methods except when the exposure time is very low (about 30 – 50ms). For this range of exposure time, the proposed method exploiting additional 3D information is able to provide better results. The MS-VST method would also potentially produce better results using a more adapted wavelet basis. For each image, the square root of the mean squared error is displayed. A reference image is defined as the average of the images displayed on the last row corresponding to an exposure time of 500ms. In order to compare the denoising results with different exposure times to this reference image, the histograms have to be aligned. A linear relationship is assumed between the intensity of each image and the reference image. Once the parameters have been estimated using a linear regression, the intensity can be corrected and the mean squared error computed. This procedure does not take into account possible motions between frames. However excepted for  $t = 50$ ms, the images were aligned. Moreover, motion compensation would imply the interpolation of noisy data and could therefore introduce potential artifacts. The mean squared

TABLE II

INFLUENCE OF THE VARIANCE STABILIZATION TRANSFORM AND OF THE USE OF TEMPORAL INFORMATION ON THE ERROR. THREE NORMS ARE USED TO MEASURE THE PERFORMANCE OF THE DENOISING METHOD. THE MEAN AND STANDARD DEVIATION WITH RESPECT TO TIME ARE REPORTED. THE COMPUTATION TIMES  $t_e$  FOR EACH EXPERIMENT IS ALSO GIVEN FOR THE NOISY SEQUENCE; 3D+TIME - GAUSSIAN AND POISSON NOISE (A) ; 3D+TIME - GAUSSIAN NOISE (B) ; 3D - POISSON AND GAUSSIAN NOISE (C).

Sequences	$L_\infty$		$L_1$		$L_2$		$t_e$
	mean	std	mean	std	mean	std	
Noisy	62.67	4.21	4.39	$6 \cdot 10^{-3}$	35.0	$12 \cdot 10^{-3}$	
A	38.35	2.87	1.56	$16 \cdot 10^{-3}$	2.94	$28 \cdot 10^{-3}$	65 min
B	53.10	5.83	1.96	$17 \cdot 10^{-3}$	3.78	$25 \cdot 10^{-3}$	55 min
C	37.98	2.44	1.65	$14 \cdot 10^{-3}$	3.01	$24 \cdot 10^{-3}$	28 min

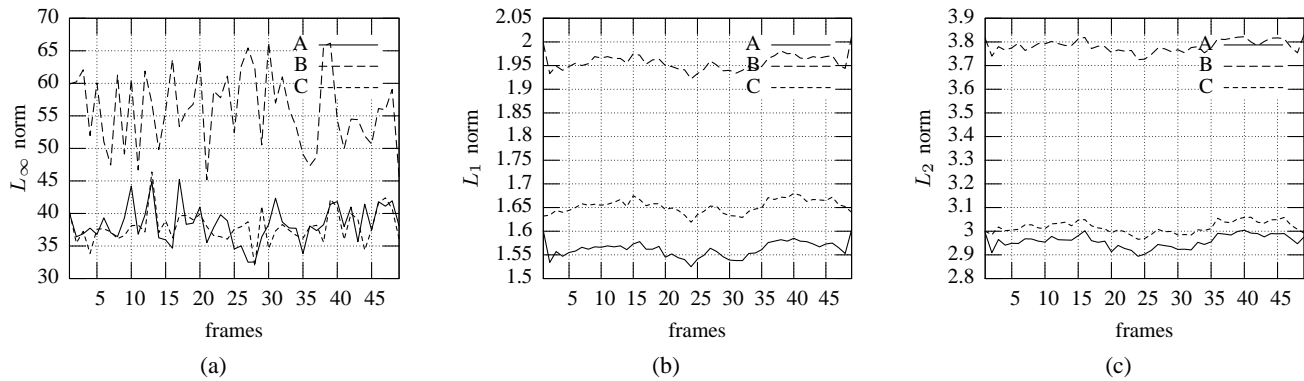


Fig. 5. Influence of the variance stabilization transform and the adjacent temporal volumes on the signal-to-noise ratios. (See text)

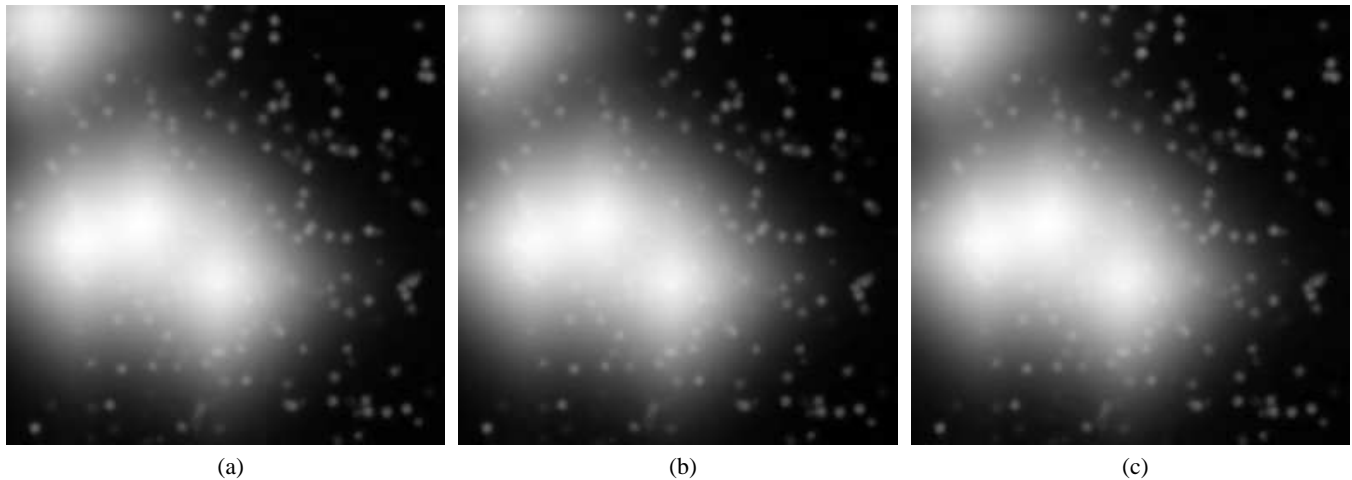


Fig. 6. XY slices #5 at time  $t = 25$  of the denoised synthetic image sequence corresponding to experiments A , B and C, respectively in (a), (b) and (c) (logarithmic scale).

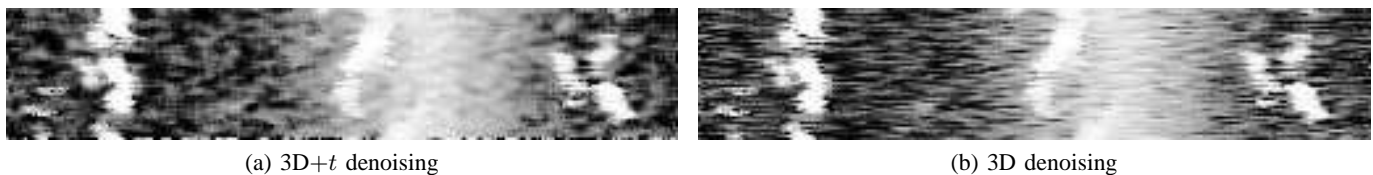


Fig. 7. YT slice #5 at  $x = 250$  of the denoised synthetic image sequence corresponding to experiments A and C, respectively in (a) and (b) after histogram equalization. More flickering effects are visible when the volumes are independently processed.

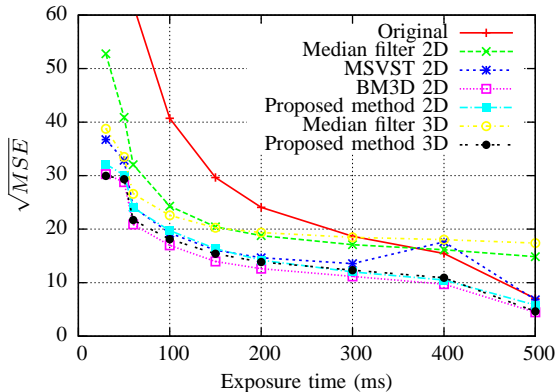


Fig. 9. Square root of the mean squared error is plotted against the exposure time showing the improvement of the filtering in the case of a fixed sample (see Fig. 8).

error values of all the experiments are summarized in Fig. 9. This experiment allows to make a direct link between the image quality and the exposure time. However due to the normalization procedure and the possible motions, the results have to be interpreted carefully and depend as well on the image content. Finally, note that exploiting temporal information would increase even more the quality of the images.

### C. 2D Space-time denoising of a synthetic image sequence

In order to compare the proposed method to another 2D+time denoising procedure, we have simulated a 2D image sequence having the same photometric characteristics than the original image used in the previous experiment with an exposure time of 500ms. Approximately 300 spots were detected and re-drawn on an estimated background profile. A Gaussian random walk was then applied to the spot positions. Noise has been generated using the same parameters than those estimated on the original image and the global intensity of the image has been varied by a factor 1, 3/4 and 1/2. The obtained image sequences have been then denoised using the multi-frame fast Haar wavelet denoising approach proposed in [63] (using 3 frames and 1 cycle spinning) and using the proposed method (using  $3 \times 3$  patches and 5 iterations). In both case, the noise parameters are the same than in the noise generation step. Table III contains the associated mean square errors. On this dataset, the proposed method performs slightly better than the method proposed in [63]. However, adjusting the parameters could potentially improve the first method.

### D. Real 3D+time image sequence

In this section, we evaluate the proposed denoising method on a real 3D+time image sequence composed

intensity	original	[63]	proposed method
1.00	14.85	3.96	3.11
0.75	18.21	4.60	3.66
0.50	32.82	14.3	13.8

TABLE III

SQUARE ROOT OF THE MEAN SQUARE ERROR FOR SEVERAL INTENSITY LEVELS USING A SIMULATED 2D IMAGE SEQUENCE HAVING THE SAME PHOTOMETRIC PROPERTIES THAN THE REFERENCE IMAGE IN FIG. 8. THE RESULTS OF THE MULTI-FRAME FAST HAAR WAVELET DENOISING [63] AND OF OUR METHOD ARE REPORTED.

of 50 volumes of  $696 \times 520 \times 6$  voxels. The slice #3 extracted at time  $t = 20$  is displayed in Fig. 11(a). This sequence has been acquired using a “fast” 4D wide-field microscope. The biological sample is a chimeric protein construct between GFP and Rab6A (GFP-RAB6A) a member of the Rab-GTPase proteins reversibly bounded to specific membranes within the living cell. At the steady state, this protein is associated to the Golgi apparatus as well as to rapidly moving transport intermediates and is present in the cytosol. Cellular dynamics of Rab6A is influenced by at least three distinct phenomena: i) lateral diffusion dictated by lipid movement within a continuum of membranes ; ii) continuous exchange between cytosolic and membrane bound pools ; iii) directional motion on membrane transport intermediates. In the sequence, the Rab6A proteins appear as dark spots when associated to small moving vesicles inside the living cell. The large dark stable structure corresponds to the Golgi apparatus while the background of the cell reveals its presence in the cytosol.

The estimation of the parameters of the generalized Anscombe transform is illustrated in Fig. 10. The regression line has been estimated and we found  $\widehat{\text{Var}}[Z_i] = 0.447 \widehat{\mathbb{E}}[Z_i] - 33.15$ . As shown in Fig. 10(b), once stabilized, the noise variance is 1.01. The results obtained with our denoising method ( $5 \times 5 \times 5$  patches) are reported in Fig. 11(b). Again, we can notice that the noise has been strongly reduced and that fine details like fluorescent particles are well preserved. The computation time for the whole volume sequence is about 80min using a standard C++ implementation. Experiments on numerous volume sequences confirm the ability of the proposed method to preserve space-time discontinuities.

### E. Combining denoising and deconvolution

Wide-field deconvolution microscopy has been widely used this last twenty years in cell biology [64], [65] as a regular tool for monitoring the living cell activity at high spatial and temporal resolution. Compared to

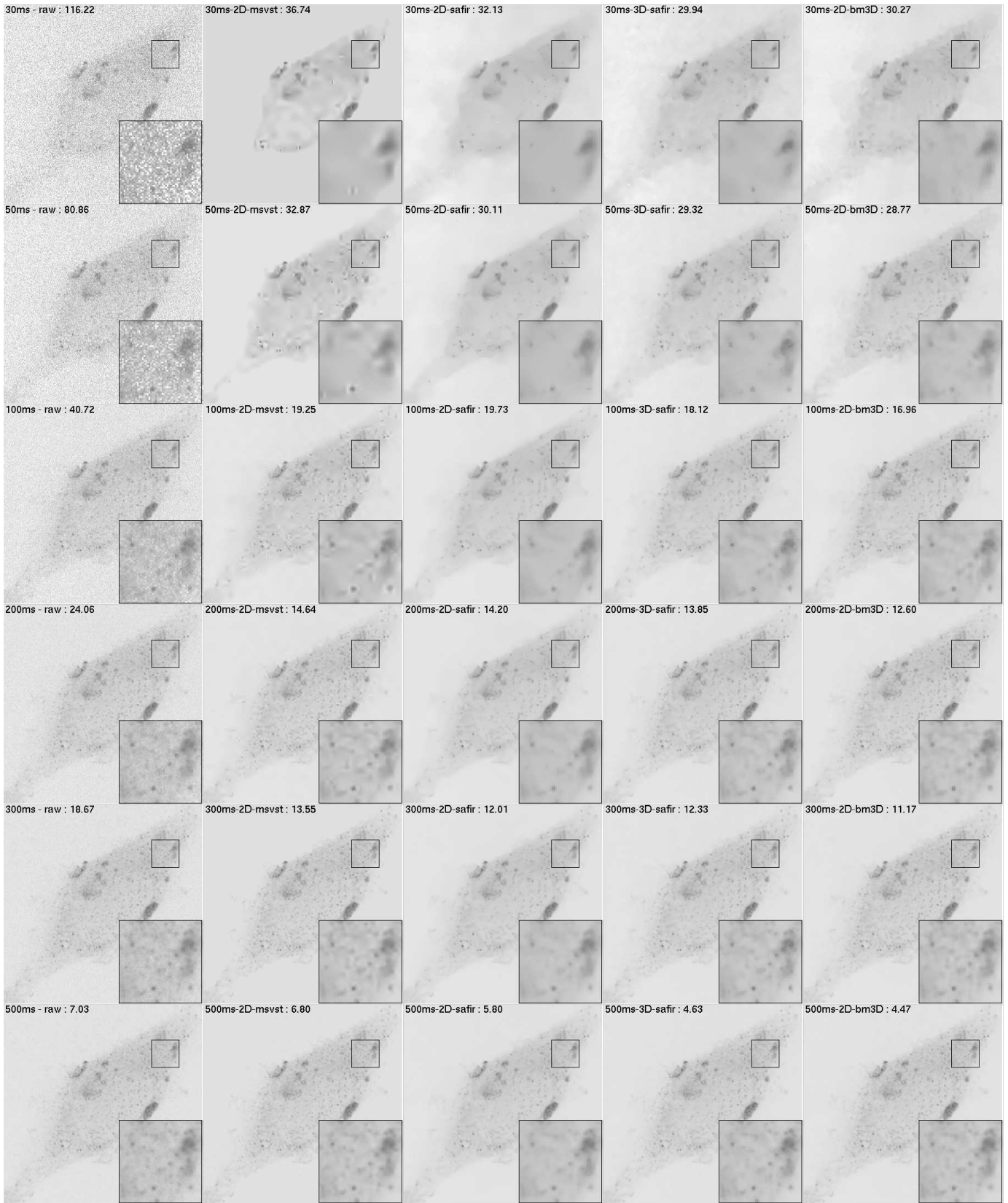


Fig. 8. Experiments on a fixed HeLa cell tagged with GFP-Rab6 acquired in spinning disk microscopy. The first column contains a 2D slice of the original 3D images taken with exposure times ranging from 30ms to 500ms. The corresponding PPSNR is increasing logarithmically with the exposure time from 25.83dB to 38.41dB. The second and third columns represent the corresponding denoising results obtained respectively with the multi-scale variance stabilization method [62] using an anisotropic wavelet basis, the proposed method in 2D and 3D, and the BM3D method [59] using the proposed variance stabilization. The numbers indicated correspond to the  $\sqrt{\text{MSE}}$  computed using the mean of the images obtained for 500ms of exposure time (last row) as a reference image.

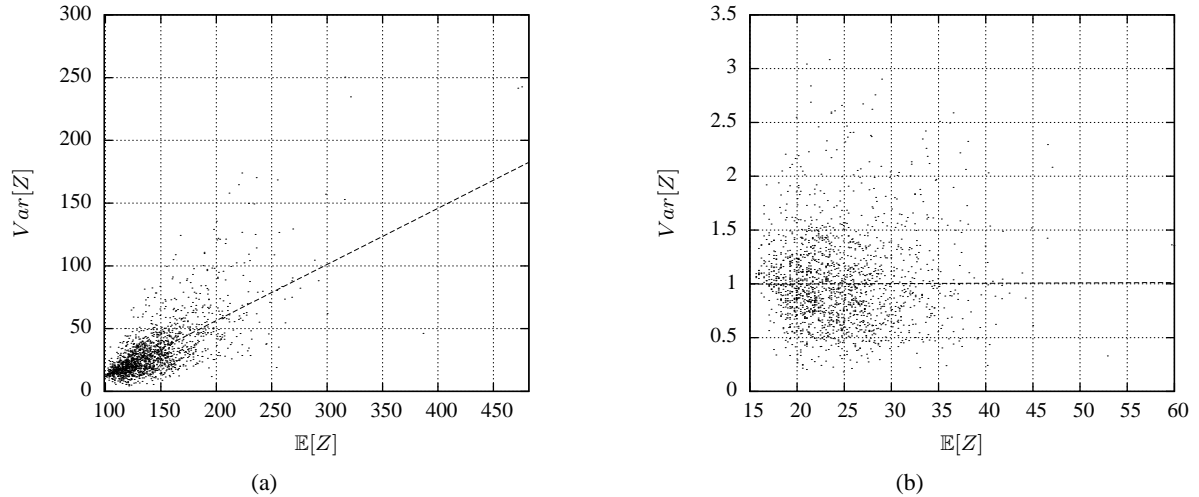


Fig. 10. Noise variance stabilization for the real image sequence shown in Fig. 11a. Estimation of the local mean  $\mathbb{E}[Z_i]$  and local variance  $\text{Var}[Z_i]$  (a) before stabilization and (b) after stabilization.

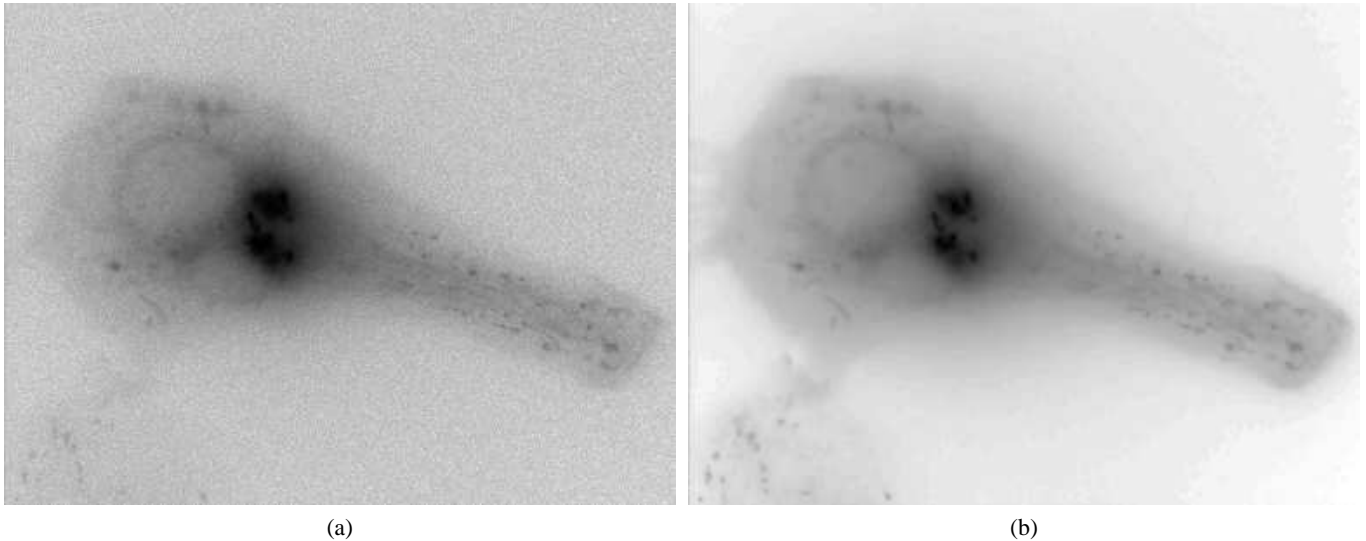


Fig. 11. Denoising of a wide-field microscopy image sequence of 50 volumes of size  $696 \times 520 \times 6$  voxels. The slice #3 of the original volume at time  $t = 20$  is displayed in (a) and the corresponding denoised volume is shown in (b) (logarithmic scale). As a result of photo-bleaching, the PPSNR decreases along time from 37.29dB to 36.75dB.

confocal like microscopy, it has the advantage to be faster, because of the wide-field illumination, and more efficient thanks to the absence of pinhole to reject photons and the highest quantum efficiency of detectors. Out-of-focus information is used and computationally reassigned to its original location, therefore increasing contrast and signal-to-noise ratio. It is known that the two main limitations of photonic microscopy are i) spatial resolution due to diffraction limit of optics and ii) the number of photons reaching the detector to statistically form the diffraction limited image. In modern living cell microscopy, the number of photons is decreased as much as possible in order to reduce the radiation dose

on the sample to keep the cell alive and to increase the acquisition frame rate. The main limitation resides in the limited number of emitted photons reaching the detector to form an image. In addition, deconvolution algorithm efficiency is sensitive to the image signal-to-noise ratio (SNR). The smaller the SNR is the less the algorithms are capable to restore the relevant signal from the noise, up to not being able to make the difference between noise and signal, resulting in artifacts.

In this section, we propose to combine the proposed denoising approach with an iterative constrained Gold-Meinel deconvolution method [66] using a fixed biological sample. Although this deconvolution method does not

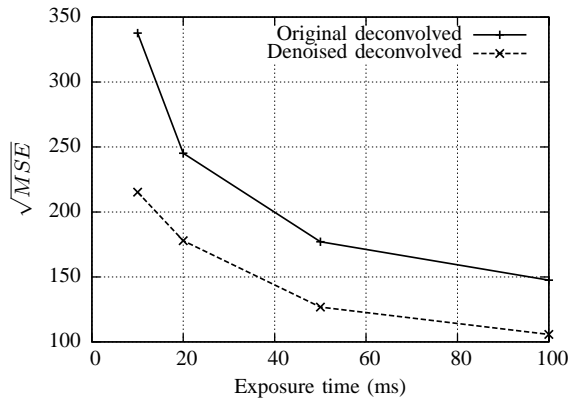


Fig. 14. The square root of the mean squared errors is plotted against the exposure times in the case of a fixed sample shown in Fig. 12. The Gold-Meinell deconvolution algorithm is applied respectively to the original and denoised images.

represent the state of the art, it shows a good robustness to the inaccuracy of the point spread function. Moreover, it is widely used and therefore the combination with the proposed denoising method is of interest.

In the same fashion than in Section IV-B, we propose to compare stacks acquired with several exposure times ranging from 10ms to 100ms to a reference image acquired with an exposure time of 200ms. Figure 12 shows the maximum intensity projection of the results. The intensity of original image shown in the first row ranges from 96 – 260 gray levels for the image acquired at 10ms of exposure time to 124 – 3315 gray levels for the image acquired at 200ms of exposure time. Figure 13 shows a zoomed area of an optical section and intensity profiles along a microtubule (polymers of  $\alpha$ - and  $\beta$ -tubulin dimers which are one of the components of the cytoskeleton). This illustrates that fine details are preserved and that the noise level is strongly reduced. Finally, mean squared errors, computed on normalized images and displayed in Fig. 14, confirm that the deconvolution is improved if the denoising is applied beforehand.

## V. CONCLUSION

In this paper, we have first tackled the issue of modeling a 3D+time video-microscopy image sequence. We have proposed to use the generalized Anscombe transform to stabilize the variance of the Poisson and Gaussian noise. We have introduced a patch-based functional and we have shown that the fixed-point solution yields an estimator involving image patches taken in a spatially-varying neighborhood. The analysis of the bias-variance of this estimator enables to properly select, for each point of the space-time domain, the optimal bandwidth within a sequence of increasing bandwidths. Spatial and

temporal dimensions are adequately handled. The overall method involves a limited number of parameters so that we do not have to tune them in practice.

We have demonstrated that the proposed method outperforms other very competitive methods in 2D and 2D+time. Moreover, experiments on real image sequences show that the space-time discontinuities are well preserved without motion estimation. Finally, we have used the capability of the proposed algorithm to efficiently denoise 3D images in order to use it as a pre-processing step prior to deconvolution. We have illustrated the efficiency of such a combination to restore low signal-to-noise ratio images. This opens interesting perspectives for monitoring biological samples at high temporal and spatial resolution, without increasing the radiation dose. To conclude, we point out that the proposed method is not restricted to video-microscopy, but could deal with other 2D+time as well as 3D+time noisy image modalities, provided that an appropriate noise modeling is adopted. In this respect, this “breaking sensitivity barrier” approach advantageously completes “breaking resolution barrier” new optics [67].

## VI. ACKNOWLEDGMENT

The authors would like to thank Florian Luisier, Bo Zhang and Alexandro Foi for providing an implementation of their respective methods.

## REFERENCES

- [1] A. Buades, B. Coll, and J.-M. Morel, “A review of image denoising algorithms, with a new one,” *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 490–530, 2005.
- [2] V. Katkovnik, A. Foi, K. Egiazarian, and J. Astola, “From local kernel to non-local multiple model image denoising,” *IJCV*, 2009, to appear.
- [3] R. M. Willett and R. D. Nowak, “Platelets: a multiscale approach for recovering edges and surfaces in photon-limited medical imaging,” *IEEE Trans. on TMI*, vol. 22, no. 3, pp. 332–350, Mar. 2003.
- [4] F. J. Anscombe, “The transformation of Poisson, binomial and negative-binomial data,” *Biometrika*, vol. 35, pp. 246–254, 1948.
- [5] J.-L. Starck, F. Murtagh, and A. Bijaoui, *Image Processing and Data Analysis, the Multiscale Approach*. Cambridge University Press, 1998.
- [6] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, “Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data,” *IEEE Trans. on IP*, 2008.
- [7] P. Fryzlewicz and G. P. Nason, “A Haar-Fisz algorithm for Poisson intensity estimation,” *Journal of Computational and Graphical Statistics*, vol. 13, no. 3, pp. 621–638, Sept. 2004.
- [8] M. Jansen, “Multiscale poisson data smoothing,” *J. Roy. Statist. Soc. B*, vol. 68, no. 1, pp. 27–48, 2006.
- [9] J. Lee, “Speckle analysis and smoothing of synthetic aperture radar images,” *CGIP*, vol. 17, no. 1, pp. 24–32, Sept. 1981.



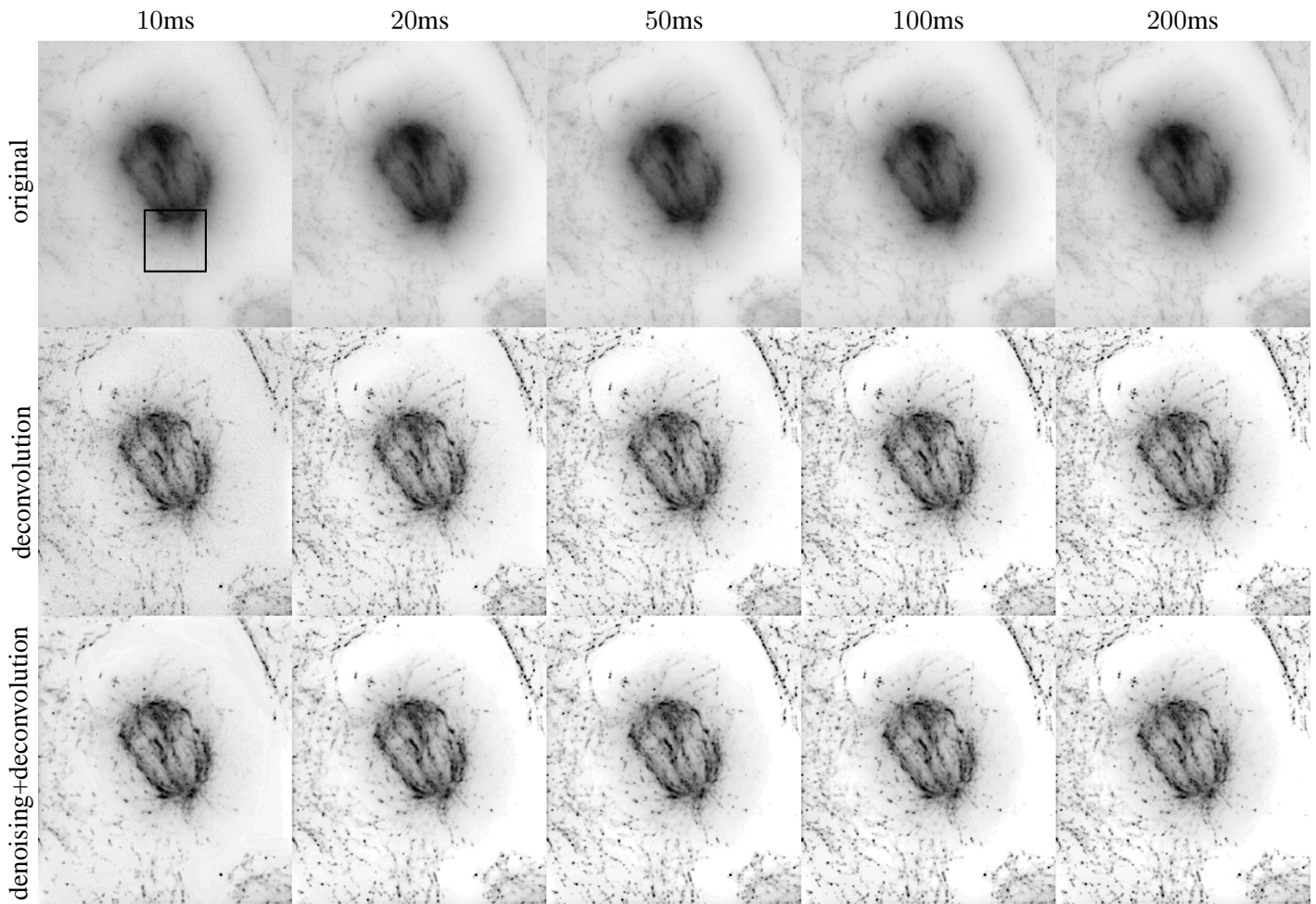


Fig. 12. A fixed HeLa cell is acquired with five increasing exposure times. The first row contains the maximum intensity projection along  $z$  direction of the  $200 \times 200 \times 36$  original images. The two last rows correspond respectively to results obtained with the Gold-Meinell deconvolution algorithm [66] and the combination with the proposed patch-based denoising.

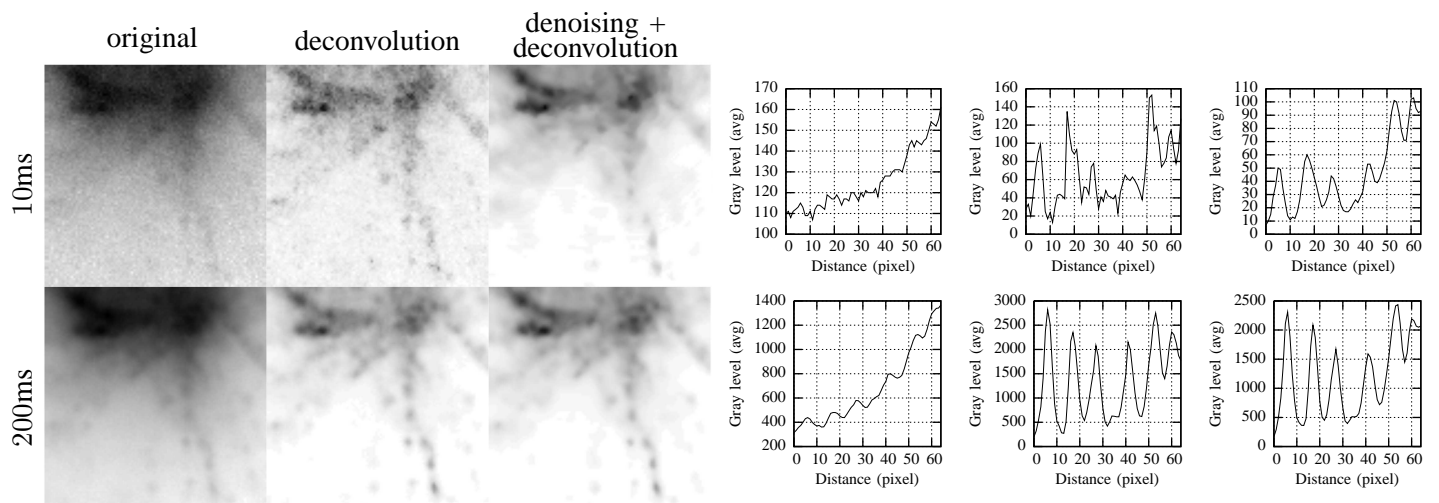


Fig. 13. Zoom of a single optical section extracted from data shown in Fig. 12 corresponding to the two extreme exposure times. The columns correspond respectively to the maximum intensity of the raw image, the results obtained with the Gold-Meinell deconvolution algorithm [66] and its combination with the proposed patch-based denoising. Plots show intensity profiles along a single microtubule for each image.

- [10] D. T. Kuan, A. A. Sawchuk, T. V. Strand, and P. Chavel, "Adaptive noise smoothing filter for images with signal-dependent noise," *IEEE Trans. on PAMI*, vol. 7, no. 2, pp. 165–177, Feb. 1985.
- [11] C. L. Chan, A. K. Katsagellos, and A. V. Sahakian, "Image sequence filtering in quantum-limited noise with application to low-dose fluoroscopy," *IEEE Trans. on TMI*, vol. 12, no. 3, pp. 610–621, Sept. 1993.
- [12] B. Zhang, M. Fadili, and J.-L. Starck, "Wavelets, ridgelets and curvelets for Poisson noise removal," *IEEE Trans. on IP*, vol. 17, no. 7, pp. 1093–1108, July 2008.
- [13] A. Santamaria Pang, T. Bildea, S. Tan, and I. Kakadiaris, "Denoising for 3-d photon-limited imaging data using nonseparable filterbanks," *IEEE Trans. on IP*, vol. 17, no. 12, pp. 2312–2323, Dec. 2008.
- [14] J. C. Brailean, R. P. Kleihorst, S. Efstratiadis, A. K. Katsagellos, and R. L. Lagendijk, "Noise reduction filters for dynamic image sequences: A review," *Proc. of the IEEE*, vol. 83, no. 9, pp. 1272–1291, Sept. 1995.
- [15] R. Dugad and N. Ahuja, "Video denoising by combining Kalman and Wiener estimates," in *Proc. of ICIP'1999*, vol. 4, Kobe, Japan, Oct. 1999, pp. 152–156.
- [16] H.-Y. Cheong, A. Tourapis, J. Llach, and J. Boyce, "Adaptive spatio-temporal filtering for video-denoising," in *Proc. of ICIP'2004*, vol. 2, Singapore, Oct. 2004, pp. 965–968.
- [17] V. Zlokolic and W. Philips, "Motion and detail adaptive denoising in video," *Proc. of SPIE – Image Processing: Algorithms and Systems III*, vol. 5298, pp. 403–412, May 2004.
- [18] A. Kuznetsov, V. P. Bindokas, J. D. Marks, and L. H. Philipson, "FRET-based voltage probes for confocal imaging: membrane potential oscillations throughout pancreatic islets," *American Journal Physiology - Cell Physiology*, vol. 289, no. 1, pp. 224–229, Mar. 2005.
- [19] N. Rajpoot, Z. Yao, and R. Wilson, "Adaptive wavelet restoration of noisy video sequences," in *Proc. of ICIP'2004*, Singapore, 2004, pp. 957–960.
- [20] F. Shi and I. W. Selesnick, "Video denoising using oriented complex wavelet transforms," in *Proc. of the IEEE ICASSP'04*, vol. 2, Montreal, Canada, May 2004, pp. 949–952.
- [21] F. Dekeyser, P. Bouthemy, and P. Pérez, "Spatio-temporal Wiener filtering of image sequences using a parametric motion model," in *Proc. of ICIP'2000*, Vancouver, Canada, Sept. 2000, pp. 208–211.
- [22] S. H. Lee and M. G. Kang, "Spatio-temporal video filtering algorithm based on 3-D anisotropic diffusion equation," in *Proc. of ICIP'1998*, vol. 3(2), Chicago, IL, Oct. 1998, pp. 447–450.
- [23] D. Uttenweiler, C. Weber, B. Jähne, R. Fink, and H. Schar, "Spatio-temporal anisotropic diffusion filtering to improve signal to noise ratios and object restoration in fluorescence microscopic image sequences," *Journal of Biomedical Optics*, vol. 8, no. 1, pp. 40–47, Jan. 2003.
- [24] S. Delpretti, F. Luisier, S. Ramani, T. Blu, and M. Unser, "Multiframe SURE-let denoising of timelapse fluorescence microscopy images," in *Proc. of IEEE ISBI'2008*, Paris, France, May 2008, pp. 149–152.
- [25] G. Motta, E. Ordentlich, I. Ramirez, G. Seroussi, and M. Weinberger, "The dude framework for continuous tone image denoising," in *Proc. of ICIP'2005*, vol. 3, 2005, pp. 345–348.
- [26] S. Awate and R. Whitacker, "Unsupervised, information theoretic, adaptive image filtering for image restoration," *IEEE Trans. on PAMI*, vol. 28, no. 3, pp. 364–376, Mar. 2006.
- [27] A. Buades, B. Coll, and J. Morel, "Nonlocal image and movie denoising," *IJCV*, vol. 76, no. 2, pp. 123–139, Feb. 2008.
- [28] A. Efros and T. Leung, "Texture synthesis by non-parametric sampling," in *Proc. of ICCV'99*, Kerkyra, Greece, Sept. 1999, pp. 1033–1038.
- [29] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based inpainting," *IEEE Trans. on IP*, vol. 13, no. 9, pp. 1200–1212, Sept. 2004.
- [30] Y. Wexler, E. Shechtman, and M. Irani, "Space-time video completion," in *Proc. of IEEE CVPR'2004*, vol. 1, Washington, 2004, pp. 120–127.
- [31] M. Mahmoudi and G. Sapiro, "Fast image and video denoising via nonlocal means of similar neighborhoods," *Signal Processing Letters*, vol. 12, no. 12, pp. 839–842, Dec. 2005.
- [32] D. Rusanovskyy, K. Dabov, and K. Egiazarian, "Moving-window varying size 3D transform-based video denoising," in *Proc. of 2nd International Workshop on Video Processing and Quality Metrics for Consumer Electronics, VPQM'06*, Jan. 2006.
- [33] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionary," *IEEE Trans. on IP*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.
- [34] J. Mairal, G. Sapiro, and M. Elad, "Learning multiscale sparse representation for image and video restoration," *Multiscale Modeling & Simulation*, vol. 7, no. 1, pp. 214–241, 2008.
- [35] M. Aharon, M. Elad, and A. Bruckstein, "The K-SVD: an algorithm for designing of overcomplete dictionaries for sparse representations," *IEEE Trans. on SP*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
- [36] J. Boulanger, C. Kervrann, and P. Bouthemy, "Space-time adaptation for patch-based image sequence restoration," *IEEE Trans. on PAMI*, vol. 29, no. 6, pp. 1096–1102, June 2007.
- [37] J. Boulanger, J.-B. Sibarita, C. Kervrann, and P. Bouthemy, "Non-parametric regression for patch-based fluorescence microscopy image sequence denoising," in *Proc. of IEEE ISBI'2008*, Paris, France, May 2008.
- [38] L. D. Brown, R. Zhang, and L. H. Zhao, "Root-unroot methods for nonparametric density estimation and poisson random-effects models," Department of Statistics University of Pennsylvania, Tech. Rep., 2001.
- [39] F. Murtagh, J. Starck, and A. Bijaoui, "Image restoration with noise suppression using a multiresolution support," *Astronomy and Astrophysics*, vol. 112, pp. 197–189, 1995.
- [40] B. Zhang, J. Fadili, J.-L. Starck, and J.-C. Olivo-Marin, "Multiscale variance stabilizing transform for mixed-poisson-gaussian processes and its applications in bioimaging," in *Proc. of ICIP'2007*, 2007.
- [41] J.-L. Starck and F. Murtagh, "Automatic noise estimation from the multiresolution support," *Publications of the Astronomical Society of the Pacific*, vol. 110, pp. 193–199, 1998.
- [42] J. Boulanger, C. Kervrann, and P. Bouthemy, *Biophotonics for Life Sciences and Medicine*. Fontis Media SA, Lausanne, Switzerland, 2006, ch. An adaptive statistical method for 4D-fluorescence image sequence denoising with spatio-temporal discontinuities preserving, pp. 97–113.
- [43] S. Ramani, C. Vonesch, and M. Unser, "Deconvolution of 3D micrographes with automatic risk minimization," in *Proc. of IEEE ISBI'2008*, Paris, France, 2008, pp. 732–735.
- [44] T. Gasser, L. Sroka, and C. Jennen-Steinmetz, "Residual variance and residual pattern in nonlinear regression," *Biometrika*, vol. 73, pp. 625–633, 1986.
- [45] P. J. Huber, *Robust Statistics*. John Wiley and Sons, New York, NY, 1981.
- [46] P. Rousseeuw and A. Leroy, *Robust Regression and Outlier Detection*. New-York: Wiley, 1987.
- [47] D. B. Murphy, *Fundamentals of light microscopy and electronic imaging*. Wiley, 2001.
- [48] M. Newberry, "Signal-to-noise considerations for sky-subtracted ccd data," *Publications of the Astronomical Society of the Pacific*, vol. 103, pp. 122–130, Jan. 1991.



- [49] S. Kindermann, S. Osher, and P. Jones, “Deblurring and denoising of images by nonlocal functionals,” *Multiscale Modeling & Simulation*, vol. 4, no. 4, pp. 1091–1115, 2005.
- [50] T. Brox, O. Kleinschmidt, and D. Cremers, “Efficient nonlocal means for denoising of textural patterns,” *IEEE Trans. on IP*, 2008.
- [51] N. Azzabou, N. Paragios, and F. Guichard, “Variable bandwidth image denoising using image-based noise models,” in *Proc. of IEEE CVPR’2007*, minneapolis, June 2007, pp. 1–7.
- [52] G. Gilboa and S. Osher, “Nonlocal linear image regularization and supervised segmentation,” *Multiscale Modeling & Simulation*, vol. 6, no. 2, pp. 595–630, 2007.
- [53] D. Tschumperlé and L. Brun, “Image denoising and registration by pde’s on the space of patches,” in *Int. Workshop on Local and Non-Local Approx. in Im. Proc. (LNLA’08)*, Lausanne, Switzerland, Aug. 2008.
- [54] L. Stankovic, “Performance analysis of the adaptive algorithm for bias-to-variance tradeoff,” *IEEE Trans. on SP*, vol. 52, no. 5, pp. 1228–1234, May 2004.
- [55] O. Lepski, “Asymptotically minimax adaptive estimation 1: upper bounds,” *Theory of Probability and Applications*, vol. 36, no. 4, pp. 654–659, 1991.
- [56] D. Belomestny and V. Spokoiny, “Spatial aggregation of local likelihood estimates with applications to classification,” *Ann. Statist.*, vol. 25, pp. 2287–2311, 2007.
- [57] C. Kervrann and J. Boulanger, “Optimal spatial adaptation for patch-based image denoising,” *IEEE Trans. on IP*, vol. 15, no. 10, pp. 2866–2878, 2006.
- [58] —, “Local adaptivity to variable smoothness for exemplar-based image denoising and representation,” *IJCV*, vol. 79, no. 1, pp. 45–69, Aug. 2008.
- [59] K. Dabov, V. Foi, A. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3D transform-domain collaborative filtering,” *IEEE Trans. on IP*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [60] C. Kervrann, J. Boulanger, and P. Coupé, “Bayesian non-local means filter, image redundancy and adaptive dictionaries for noise removal,” in *SSVM’07*, Ischia, May 2007.
- [61] B. Zhang, M. Fadili, and J.-L. Starck, “Fast poisson noise removal by biorthogonal haar domain hypothesis testing,” *Statistical Methodology*, vol. 5, no. 4, pp. 387–396, 2008.
- [62] —, “Multi-scale variance stabilizing transform for multi-dimensional Poisson count image denoising,” in *Proc. of IEEE ICASSP’2006*, Toulouse, France, May 2006.
- [63] L. Luisier, C. Vonesch, T. Blu, and M. Unser, “Fast haar-wavelet denoising of multidimensional fluorescence microscopy data,” in *Proc. of IEEE ISBI’2008*, vol. 1, Boston MA, USA, June 2008.
- [64] D. Agard, “Optical sectioning microscopy: cellular architecture in three dimensions,” *Annual Review of Biophysics and Bioengineering*, vol. 13, pp. 191–219, June 1984.
- [65] J.-B. Sibarita, *Microscopy Techniques*, ser. Advances in Biochemical Engineering/Biotechnology. Berlin / Heidelberg: Springer Verlag, 2005, vol. 95, ch. Deconvolution Microscopy, pp. 201–244.
- [66] J.-B. Sibarita, H. Magnin, and J. R. De Mey, “Ultra-fast 4D microscopy and high throughput distributed deconvolution,” in *Proc. of IEEE ISBI’2002*, Washington, DC, June 2002, pp. 769–772.
- [67] S. Hell, “Far-field optical nanoscopy,” *Science*, vol. 316, no. 5828, pp. 1153–1158, May 2007.