FAST HAAR-WAVELET DENOISING OF MULTIDIMENSIONAL FLUORESCENCE MICROSCOPY DATA

Florian Luisier[1], *Cédric Vonesch*[1], *Thierry Blu*[2] *and Michael Unser*[1]

[1]Biomedical Imaging Group, Ecole Polytechnique Fédérale de Lausanne, Switzerland
[2]Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong

ABSTRACT

We propose a novel denoising algorithm to reduce the Poisson noise that is typically dominant in fluorescence microscopy data. To process large datasets at a low computational cost, we use the unnormalized Haar wavelet transform. Thanks to some of its appealing properties, independent unbiased MSE estimates can be derived for each subband. Based on these Poisson unbiased MSE estimates, we then optimize linearly parametrized interscale thresholding. Correlations between adjacent images of the multidimensional data are accounted for through a sliding window approach. Experiments on simulated and real data show that the proposed solution is qualitatively similar to a state-of-the-art multiscale method, while being orders of magnitude faster.

Index Terms— Poisson noise, Fluorescence, Haar Wavelet, MSE estimation

1. INTRODUCTION

Fluorescence microscopy has gained an increasing interest in the last few decades thanks to its ability to identify tagged molecules in a given specimen. One of its main physical limitation stems from the random nature of photon emission and detection [1]. Especially under low illumination conditions (short time exposure or phototoxicity constraints), these random variations are well described by a Poisson process. Denoising then becomes an indispensable step prior to the visualization and analysis of fluorescence microscopy data. The multiresolution techniques are particularly well suited for such denoising tasks because they can yield a sparse representation of the data.

Estimating the intensity of a Poisson process is at the core of many works in statistics and, quite recently, in signal processing too. There are globally two main strategies to deal with Poisson statistics. The first consists in "gaussianizing" the Poisson process by applying a variance-stabilizing transform (VST), such as the Anscombe [2] or Haar-Fisz [3] transforms. Then, any denoiser designed for additive Gaussian white noise (AGWN) can be used [4, 5]. However, these standard VST-based method are usually not very efficient for estimating low intensities. The second strategy is the direct handling of Poisson statistics, often in a Bayesian framework [6, 7, 8]. While such approaches are robust, they require more complex estimators than in the AGWN case, due to the signal-dependent aspect of the Poisson noise. Note that an hybrid approach combining VST, hypothesis testing and sparsity-driven reconstruction with the latest multiresolution transforms has been recently proposed [9].

In this paper, we aim at devising a fast algorithm to efficiently denoise multidimensional fluorescence microscopy data. We thus propose to use the unnormalized Haar wavelet transform to process large datasets at a light computational load. This orthogonal transform has the appealing property of preserving Poisson statistics in its lowpass channel. In a non-Bayesian framework, it is then possible to derive an unbiased estimator of the mean-squared error (MSE) in the wavelet domain. Based on this Poisson unbiased MSE estimate, we optimize a subband-dependent thresholding function. The use of a simple non-redundant transform requires the development of a more sophisticated thresholding than the standard soft- or hard-threshold. In particular, we use a *linear expansion of thresholds* (LET) [10] which incorporates the interscale dependencies. For a given highpass subband, interscale predictors are extracted from the corresponding lowpass subband. The latter is also used to locally adapt the thresholds to the signal-dependent noise variance. Finally, to account for the strong correlations between adjacents frames (or slices) of the multidimensional data, we opt for a sliding-window approach (i.e. neighboring frames as predictors of the current frame), instead of time- and memory-consuming 3D transforms.

Thanks to the LET construction and the quadratic form of the Poisson unbiased MSE estimate, a multidimensional interscale thresholding involving several parameters can be quickly optimized through the resolution of a linear system of equations.

2. METHOD

2.1. Context

We denote a sequence $\mathbf{x}$ of C N-pixel images as a $C \times N$ matrix whose columns $\mathbf{x}_n$ are the values of each pixel across the whole sequence, i.e.

$$\mathbf{x} = [\mathbf{x}_1\ \mathbf{x}_2 \dots \mathbf{x}_N], \text{ where } \mathbf{x}_n = [x_{n,1}\ x_{n,2} \dots x_{n,C}]^{\mathrm{T}}$$

The pixels $y_{n,c}$ of the observed noisy sequence $\mathbf{y}$ are assumed to follow independent Poisson laws of underlying intensities $x_{n,c}$, i.e. (when dropping the indices n, c)

$$y \sim \mathcal{P}(x) \tag{1}$$

Our aim is then to find the best—in the minimum mean-squared error (MSE) sense— estimate $\hat{\mathbf{x}}$ of $\mathbf{x}$, given the noisy observation $\mathbf{y}$ only, i.e. we want to minimize:

$$\text{MSE} = \frac{1}{CN}\mathrm{Tr}\left\{(\hat{\mathbf{x}} - \mathbf{x})(\hat{\mathbf{x}} - \mathbf{x})^{\mathrm{T}}\right\} \tag{2}$$

This work was supported by the Center for Biomedical Imaging (CIBM) of the Geneva - Lausanne Universities and the EPFL, the foundations Leenaards and Louis-Jeantet, the Swiss National Science Foundation under grant 200020-109415, as well as by a Direct Grant #2050420 from the Hong Kong Research Grants Council.

2.2. Haar-Domain Unbiased MSE Estimate

In this work, we propose to denoise the data in the *unnormalized Haar* wavelet-domain. At each decomposition stage $j = 1 \dots J$, the scaling coefficients of the noise-free (resp. noisy) sequence are denoted by $\boldsymbol{\sigma}^j$ (resp. $\mathbf{s}^j$), while the associated wavelet coefficients are denoted by $\boldsymbol{\delta}^j$ (resp. $\mathbf{d}^j$). The filterbank implementation of the unnormalized Haar wavelet transform is given in Fig. 1.

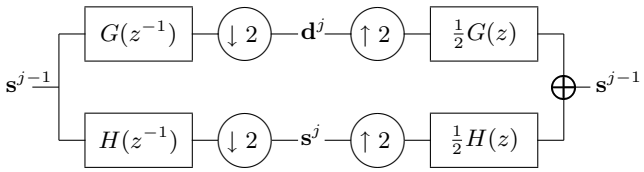

Fig. 1. One decomposition stage of the unnormalized Haar wavelet filterbank: $H(z) = 1 + z$ and $G(z) = 1 - z$ are respectively the scaling and wavelet filters. As initialization, $\mathbf{s}^0 = \mathbf{y}$.

The unnormalized Haar wavelet transform has two main advantages over more sophisticated wavelets and redundant transforms:

1. The scaling coefficients of an input vector of independent Poisson random variables are also independent Poisson random variables, i.e. (when dropping the indices n, c and the superscript j)
$$s \sim \mathcal{P}(\sigma)$$
2. It is an *orthogonal* transform. Therefore, the MSEs inside each subband can be minimized independently, while ensuring the minimization of the global MSE.

While simple thresholding rules (such as hard- or soft-threshold) are already efficient when applied in highly redundant representations, more sophisticated denoising functions must be considered in the unnormalized Haar wavelet domain. We thus propose to construct an estimate $\hat{\boldsymbol{\delta}}^j$ of the noise-free wavelet coefficients $\boldsymbol{\delta}^j$, that takes into account both the corresponding noisy wavelet coefficients $\mathbf{d}^j$ and the scaling coefficients at the same scale $\mathbf{s}^j$:

$$\hat{\boldsymbol{\delta}}^j = \boldsymbol{\theta}^j(\mathbf{d}^j, \mathbf{s}^j) \tag{3}$$

Since the noise-free wavelet coefficients $\boldsymbol{\delta}^j$ are not accessible, we cannot minimize the actual subband MSE defined by:

$$\text{MSE}_j = \frac{1}{CN_j}\text{Tr}\left\{(\boldsymbol{\theta}^j(\mathbf{d}^j, \mathbf{s}^j) - \boldsymbol{\delta}^j)(\boldsymbol{\theta}^j(\mathbf{d}^j, \mathbf{s}^j) - \boldsymbol{\delta}^j)^{\mathrm{T}}\right\}$$

where N_j is the number of multidimensional pixels in the subband at scale j.

To overcome this difficulty, we devise a statistical unbiased estimator of the actual subband MSE that only involves the observed noisy data.

Theorem 1. *Let $\boldsymbol{\theta}(\mathbf{d}, \mathbf{s}) = \boldsymbol{\theta}^j(\mathbf{d}^j, \mathbf{s}^j)$ be an estimate of the noise-free wavelet coefficients $\boldsymbol{\delta} = \boldsymbol{\delta}^j$ and let $\mathbf{e}_{c,n}$ denote a $C \times N$ matrix filled with zeros, except at the position (c, n) which is set to one. Define $\boldsymbol{\theta}^+(\mathbf{d}, \mathbf{s})$ and $\boldsymbol{\theta}^-(\mathbf{d}, \mathbf{s})$ by*

$$\begin{cases} \boldsymbol{\theta}^+(\mathbf{d}, \mathbf{s}) = [\theta_{c,n}(\mathbf{d} + \mathbf{e}_{c,n}, \mathbf{s} - \mathbf{e}_{c,n})]_{1 \le c \le C, 1 \le n \le N_j}\,, \\ \boldsymbol{\theta}^-(\mathbf{d}, \mathbf{s}) = [\theta_{c,n}(\mathbf{d} - \mathbf{e}_{c,n}, \mathbf{s} - \mathbf{e}_{c,n})]_{1 \le c \le C, 1 \le n \le N_j}\,. \end{cases} \tag{4}$$

Then the random variable

$$\begin{aligned} \epsilon_j \;=\; & \frac{1}{CN_j}\Big(\text{Tr}\left\{\boldsymbol{\theta}(\mathbf{d}, \mathbf{s})\boldsymbol{\theta}(\mathbf{d}, \mathbf{s})^{\mathrm{T}} + \mathbf{d}\mathbf{d}^{\mathrm{T}} - \mathbf{1}\mathbf{s}^{\mathrm{T}}\right\} \\ & -\text{Tr}\left\{\mathbf{d}\big(\boldsymbol{\theta}^-(\mathbf{d}, \mathbf{s}) + \boldsymbol{\theta}^+(\mathbf{d}, \mathbf{s})\big)^{\mathrm{T}}\right\} \\ & -\text{Tr}\left\{\mathbf{s}\big(\boldsymbol{\theta}^-(\mathbf{d}, \mathbf{s}) - \boldsymbol{\theta}^+(\mathbf{d}, \mathbf{s})\big)^{\mathrm{T}}\right\}\Big) \end{aligned} \tag{5}$$

is an unbiased *estimate of the MSE for the subband under consideration, i.e., $\mathscr{E}\{\epsilon_j\} = \mathscr{E}\{MSE_j\}$.*

Proof. The proof of Theorem 1 is mainly based on relation (5.1) in [11] . □

The accuracy of the proposed Poisson unbiased MSE estimate improves with the data size. Therefore, (5) is close to the actual MSE in fluorescence microscopy, due to the high number of acquired pixels.

2.3. Linear Expansion of Thresholds for Fast Denoising

To take advantage of the quadratic form of the unbiased MSE estimate (5), we built our wavelet estimator as a *linear expansion of thresholds* [10], i.e.

$$\begin{aligned} \boldsymbol{\theta}(\mathbf{d}, \mathbf{s}) \;&=\; \sum_{k=1}^{K} \mathbf{a}_k^{\mathrm{T}} \boldsymbol{\theta}_k(\mathbf{d}, \mathbf{s}) \\ &=\; \underbrace{[\mathbf{a}_1^{\mathrm{T}}, \mathbf{a}_2^{\mathrm{T}}, \dots, \mathbf{a}_K^{\mathrm{T}}]}_{\mathbf{A}^{\mathrm{T}}} \begin{bmatrix} \boldsymbol{\theta}_1(\mathbf{d}, \mathbf{s}) \\ \boldsymbol{\theta}_2(\mathbf{d}, \mathbf{s}) \\ \vdots \\ \boldsymbol{\theta}_K(\mathbf{d}, \mathbf{s}) \end{bmatrix} \end{aligned} \tag{6}$$

Thanks to this linear parameterization, the optimal set of $KC \times C$ parameters $\mathbf{A}$ (i.e. the minimizer of (5)) is the solution of a *linear* system of equations:

$$\mathbf{A}_{\text{opt}} = \mathbf{M}^{-1}\mathbf{B} \tag{7}$$

where

$$\begin{cases} \mathbf{M} \;=\; [\boldsymbol{\theta}_k(\mathbf{d}, \mathbf{s})\boldsymbol{\theta}_l(\mathbf{d}, \mathbf{s})^{\mathrm{T}}]_{1 \le k,l \le K} \\ \mathbf{B} \;=\; \frac{1}{2}[\mathbf{d}\big(\boldsymbol{\theta}_k^-(\mathbf{d}, \mathbf{s}) + \boldsymbol{\theta}_k^+(\mathbf{d}, \mathbf{s})\big)^{\mathrm{T}}]_{1 \le k \le K} \\ \qquad\; + \frac{1}{2}[\mathbf{s}\big(\boldsymbol{\theta}_k^-(\mathbf{d}, \mathbf{s}) - \boldsymbol{\theta}_k^+(\mathbf{d}, \mathbf{s})\big)^{\mathrm{T}}]_{1 \le k \le K} \end{cases} \tag{8}$$

2.4. Multidimensional Interscale Thresholding for Poisson Noise

In the Haar wavelet transform, there is no group delay between the lowpass $H(z)$ and highpass $G(z)$ analysis filters. Therefore, an interscale predictor $\tilde{\mathbf{d}}_n$ of $\mathbf{d}_n$ can be easily constructed from the same scale lowpass subband $\mathbf{s}$, as $\tilde{\mathbf{d}}_n = \mathbf{s}_{n-1} - \mathbf{s}_{n+1}$. By construction, the sign of $\tilde{\mathbf{d}}_n$ turns out to be consistent with those of the corresponding wavelet coefficient $\mathbf{d}_n$. Therefore, $\tilde{\mathbf{d}}_n$ can be directly integrated in the linear expansion of thresholds.

Grouping together wavelet coefficients of similar magnitudes can also bring some improvements. This grouping can be refined by taking into account the magnitude of $\tilde{\mathbf{d}}_n$; in practice, to increase the robustness toward noise, it is more efficient to use the magnitude of a smoothed version[1] $\mathbf{p}_n$ of the interscale predictor $\tilde{\mathbf{d}}_n$.

[1]This smoothed version is obtained by applying a normalized Gaussian kernel on the absolute value of $\tilde{\mathbf{d}}$.

Similarly to the multiframe denoising of additive Gaussian white noise [5], we thus propose to use a multidimensional interscale thresholding of the following form:

$$\begin{aligned}\boldsymbol{\theta}_n(\mathbf{d},\mathbf{s}) \quad = \quad & \underbrace{\gamma(\mathbf{p}_n^{\mathrm{T}}\mathbf{p}_n)\gamma(\mathbf{d}_n^{\mathrm{T}}\mathbf{d}_n)}_{\text{small predictors and small coefficients}} \mathbf{a}_1^{\mathrm{T}}\mathbf{d}_n + \\ & \underbrace{\overline{\gamma}(\mathbf{p}_n^{\mathrm{T}}\mathbf{p}_n)\gamma(\mathbf{d}_n^{\mathrm{T}}\mathbf{d}_n)}_{\text{large predictors and small coefficients}} \mathbf{a}_2^{\mathrm{T}}\mathbf{d}_n + \\ & \underbrace{\gamma(\mathbf{p}_n^{\mathrm{T}}\mathbf{p}_n)\overline{\gamma}(\mathbf{d}_n^{\mathrm{T}}\mathbf{d}_n)}_{\text{small predictors and large coefficients}} \mathbf{a}_3^{\mathrm{T}}\mathbf{d}_n + \\ & \underbrace{\overline{\gamma}(\mathbf{p}_n^{\mathrm{T}}\mathbf{p}_n)\overline{\gamma}(\mathbf{d}_n^{\mathrm{T}}\mathbf{d}_n)}_{\text{large predictors and large coefficients}} \mathbf{a}_4^{\mathrm{T}}\mathbf{d}_n + \\ & \underbrace{\gamma(\mathbf{p}_n^{\mathrm{T}}\mathbf{p}_n)\,\mathbf{a}_5^{\mathrm{T}}\tilde{\mathbf{d}}_n + \overline{\gamma}(\mathbf{p}_n^{\mathrm{T}}\mathbf{p}_n)\,\mathbf{a}_6^{\mathrm{T}}\tilde{\mathbf{d}}_n}_{\text{sign consistency enhancement}} \end{aligned} \tag{9}$$

where $\gamma(x) = \exp\left(-\frac{|x|}{2T^2}\right)$ and $\overline{\gamma}(x) = 1-\gamma(x)$ are the two complementary grouping functions.

Contrary to the Gaussian case where the noise variance is constant, the threshold T must reflect the non-stationarity of the Poisson noise. A good estimate of the local noise variance of the wavelet coefficients $\mathbf{d}_n$ is given by the magnitude of the corresponding scaling coefficient $\mathbf{s}_n$. In our experiments, we thus found that $T^2 = 6|\mathbf{1}^{\mathrm{T}}\mathbf{s}_n|$ gave the lowest MSE.

2.5. Algorithm

In practice, the data usually follow a scaled and shifted Poisson law, due to the overall gain α and offset μ of the photodetectors, i.e. $y \sim \alpha\,\mathcal{P}(x)+\mu$. These two parameters must be estimated from the noisy samples to properly rescale the data according to the Poisson model (1). For this propose, we use the procedure described in [5].

To have a fast algorithm at a low computational cost, we use a sliding window approach, which allows the parallel denoising of each image of the sequence. In particular, only $C'(\text{odd}) \leq C$ images are used to denoise the central $c = (C'+1)/2$ image. Inside a given window, the parameters $\mathbf{a}_k$ of the thresholding function (9) are thus optimized for the central image only. An overview of the whole algorithm is depicted in Fig. 2.

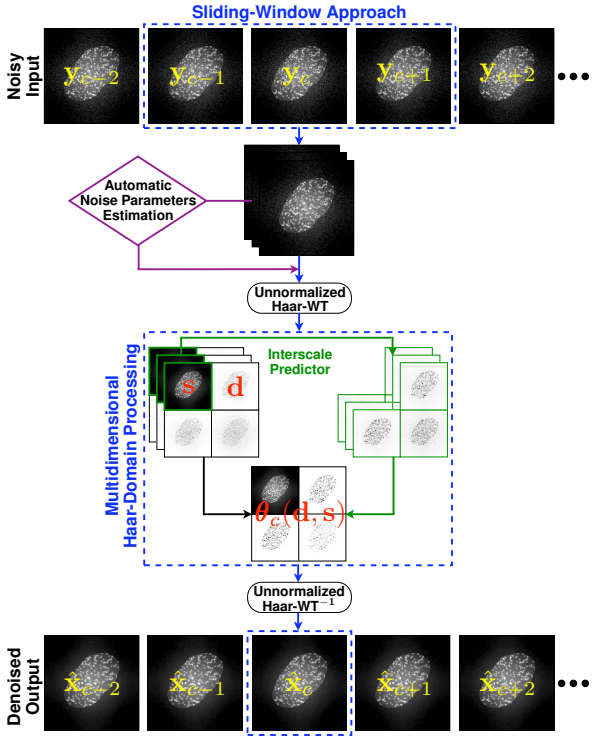

Fig. 2. Overview of the proposed denoising algorithm.

3. EXPERIMENTS

3.1. Validation on Simulated Data

To evaluate the performance of the proposed algorithm, we simulated a noisy sequence of size $512 \times 512 \times 10$ (part of it is shown in Fig. 2), with a mean intensity of 5, leading to an input signal-to-noise ratio (SNR) of 10.66 dB. The gain was set to $\alpha = 1$ and the offset to $\mu = 0$. The estimated parameters were $\alpha = 1.01$ and $\mu = 0.11$. Using these estimated parameters, we obtained an output SNR of 23.08 dB in 7.5s using $C = 3$ adjacent frames and an output SNR of 23.54 dB in 11.7s using $C = 5$ adjacent frames. We got an output SNR of 22.74 dB in 25.8s with the VST-based SURE-LET algorithm described in [5], and an output SNR of 19.85 dB in 7.5s with a 3D ($5 \times 5 \times 3$) median filter. We also applied the recent *Platelet* algorithm [2] [8] separately on each frame of the sequence: we obtained an output SNR of 22.67 dB in about 8hrs30min using 25 cyclic shifts. The computation time thus becomes a crucial point when dealing with large datasets, making the use of redundant transformations discouraging.

[2]Matlab implementation available at:
`http://people.ee.duke.edu/~willett/Research/platelets.html`

3.2. Results on Real Data

We acquired a $1024 \times 1024 \times 64$ stack of fluorescence images at the BioImaging and Optics platform (BIOP) at EPFL. We used a confocal microscope equipped with a 63X PL-APO objective. The X-Y pixel size was set to $0.09\ \mu m \times 0.09\ \mu m$ and the Z-step was $0.37 \mu m$. In addition to fibroblast cells labeled with a DiO dye (which is predominantly retained by the cell membranes), the sample contained 100-nm fluorescent microbeads acting as point sources (see Fig. 3(A)). In Fig. 3, we display the denoising results of the standard 3D median filter (B), the recent *Platelets* approach (C) and the proposed algorithm (D). Observe that our solution compares favorably with the state-of-the-art *Platelets* technique, while being orders of magnitude faster.

4. CONCLUSION

We have presented a fast and efficient algorithm to reduce Poisson noise in multidimensional fluorescence microscopy. In the unnormalized Haar wavelet domain, we have devised a prior-free unbiased estimate of the MSE on which we rely to optimize a linearly parametrized interscale thresholding. The strong correlations between adjacent images of the sequence have been incorporated through a sliding-window approach. Experiments on simulated and real data have shown that the proposed method is competitive with a state-of-the-art algorithm, while having a much lower computational cost. An ImageJ[3] plugin of the proposed algorithm will be available soon.

[3]Free software available at: `http://rsb.info.nih.gov/ij/`

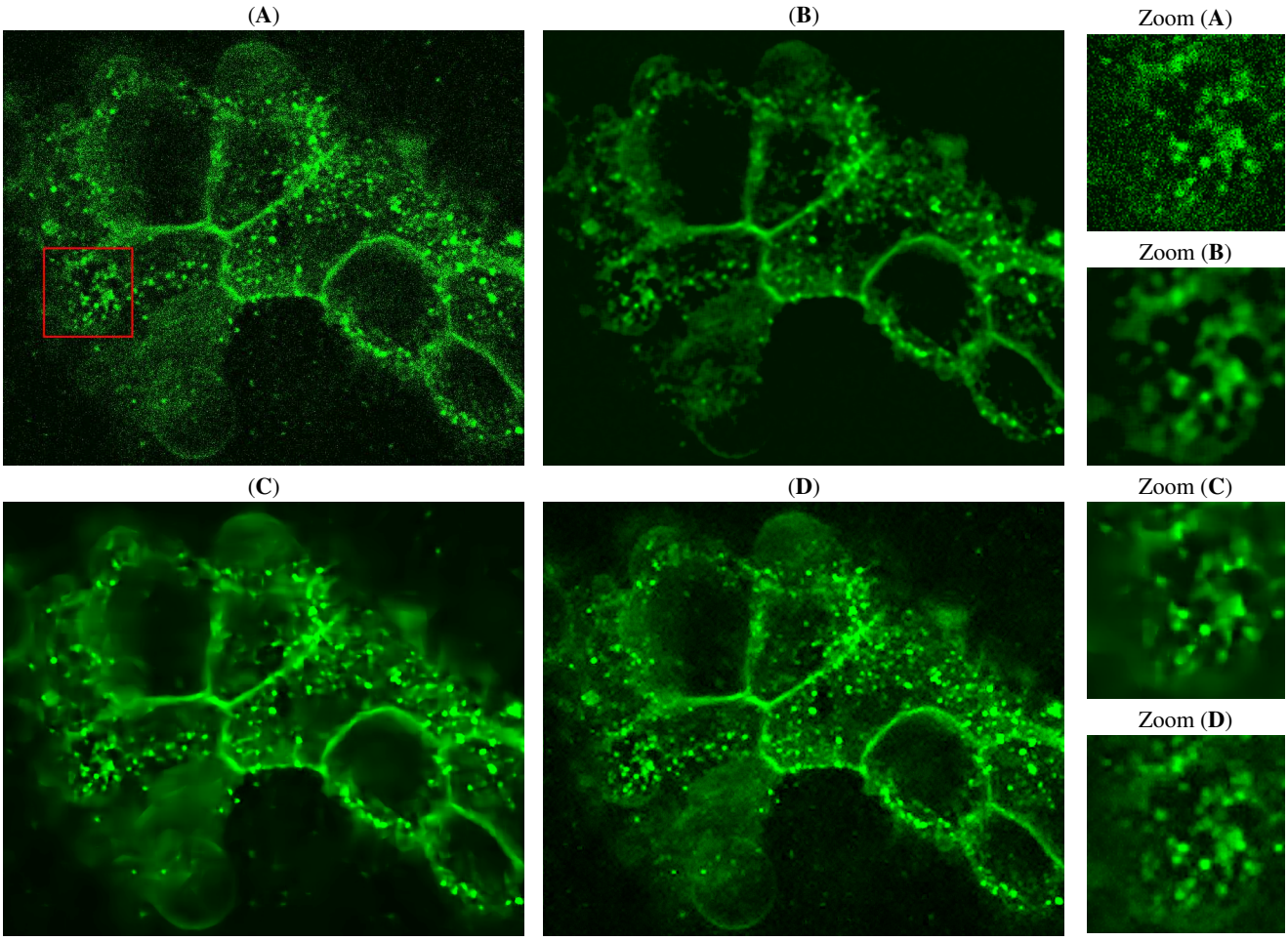

Fig. 3. Part of a particular 1024×1024 slice of the 3D fluorescence stack. (**A**) Raw slice. (**B**) $7 \times 7 \times 3$ Median filter: 8.4s. (**C**) 25 cycle-spins of *Platelets*: 42min. (**D**) The proposed algorithm using $C = 3$ adjacent slices: 3.5s.

5. REFERENCES

[1] C. Vonesch, F. Aguet, J.-L. Vonesch, and M. Unser, "The colored revolution of bioimaging," *IEEE Signal Processing Magazine*, vol. 23, no. 3, pp. 20–31, May 2006.

[2] F. J. Anscombe, "The Transformation of Poisson, Binomial and Negative-Binomial Data," *Biometrika*, vol. 35, no. 3/4, pp. 246–254, December 1948.

[3] Piotr Fryzlewicz and Guy P. Nason, "A Haar-Fisz Algorithm for Poisson Intensity Estimation," *Journal of Computational and Graphical Statistics*, vol. 13, no. 3, pp. 621–638, September 2004.

[4] J. Boulanger, J.-B. Sibarita, C. Kervrann, and P. Bouthemy, "Non-parametric regression for patch-based fluorescence microscopy image sequence denoising," *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pp. 748–751, May 2008.

[5] S. Delpretti, F. Luisier, S. Ramani, T. Blu, and M. Unser, "Multiframe SURE-LET denoising of timelapse fluorescence microscopy images," *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pp. 149–152, May 2008.

[6] K. E. Timmermann and R. D. Nowak, "Multiscale modeling and estimation of Poisson processes with application to photon-limited imaging," *IEEE Transactions on Information Theory*, vol. 45, no. 3, pp. 846–862, April 1999.

[7] E. D. Kolaczyk, "Bayesian multi-scale models for Poisson processes," *Journal of the American Statistical Association*, vol. 94, no. 447, pp. 920–933, September 1999.

[8] Rebecca M. Willett and Robert D. Nowak, "Multiscale Poisson Intensity and Density Estimation," *IEEE Transactions on Information Theory*, vol. 53, no. 9, pp. 3171–3187, September 2007.

[9] B. Zhang, J. M. Fadili, and J.-L. Starck, "Wavelets, ridgelets, and curvelets for poisson noise removal," *Image Processing, IEEE Transactions on*, vol. 17, no. 7, pp. 1093–1108, July 2008.

[10] T. Blu and F. Luisier, "The SURE-LET approach to image denoising," *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2778–2786, November 2007.

[11] H. M. Hudson, "A natural identity for exponential families with applications in multiparameter estimation," *The Annals of Statistics*, vol. 6, no. 3, pp. 473–484, May 1978.