

AN ADAPTIVE STATISTICAL METHOD FOR DENOISING 4D FLUORESCENCE IMAGE SEQUENCES WITH PRESERVATION OF SPATIO-TEMPORAL DISCONTINUITIES

J. Boulanger, C. Kervrann

IRISA - INRIA / INRA MIA
Campus Universitaire de Beaulieu
35042 Rennes Cedex - France

P. Bouthemy

IRISA - INRIA
Campus Universitaire de Beaulieu
35042 Rennes Cedex - France

ABSTRACT

We present a spatio-temporal filtering method for significantly increasing the signal-to-noise ratio in noisy fluorescence microscopic image sequences where small particles have to be tracked from frame to frame. Image sequences restoration is achieved using a spatio-temporal adaptive window approach with an appropriate on-line window geometry specification. We have applied this method to noisy synthetic and real 3D image sequences where a large number of small fluorescently labelled vesicles are moving in regions close to the Golgi apparatus. The SNR is shown to be drastically improved and the enhanced vesicles can be segmented. This novel approach can be further used for biological studies where the dynamic of small objects of interest has to be analyzed in molecular and sub-cellular bio-imaging.

1. INTRODUCTION

Fluorescence microscopy is now able to routinely produce 4D or 5D digital data sets. Both confocal microscopy and deconvolution microscopy aim at improving the spatial resolution of the image in lateral xy directions as well as in depth z . High resolution images are then relatively easy to obtain. These technologies also allow the acquisition of temporal series in eventually different wavelength (\mathcal{L}) domains. The recorded data sets have then four ($3D + t$) or five ($3D + t + \mathcal{L}$) dimensions and are crucial for dynamic understanding at sub-cellular or molecular levels. This explains their widespread in biomedicine and biophysics.

Despite the increasing resolution of images, SNR is often low due to the photon-limited nature of the images. Denoising is then an important processing step for visualization purposes and pattern recognition. Filtering especially takes an important place when dynamics have to be estimated. In several applications in biology, acquired sequences represent small fluorescently tagged particles with highly variable velocities. Moreover, trajectories are spatio-temporal line-like structures that smoothing algorithms must preserve.

However, particle motion cannot be reliably computed without a denoising pre-processing step.

In practice, microscopy image sequences are usually processed frame by frame using spatial filters. This is due to the lack of tools for 4D data even if spatio-temporal filtering, investigated for a decade, could be more appropriate. In video processing, previous commonly-used methods use temporal or spatio-temporal information, and can use motion compensation [2]. Most of spatio-temporal filters rely on an extension of well-known principles developed for 2D still image denoising, such as anisotropic diffusion [5, 6] and wavelet shrinkage [7].

In this paper we also address the sequence denoising problem and use the framework described for still images in [8]. Unlike anisotropic diffusion [1] and bilateral filtering [9], it is a data-driven estimation approach and a pointwise scale selection is obtained by locally estimating the optimal spatio-temporal filtering window. The properties of such an estimator are well known in statistics for adaptive smoothing [10, 11, 3]. Additionally, it requires no explicit *a priori* knowledge on the unknown image and provides a confidence level attached to the each pixel. The proposed method is not based on energy minimization [4] and its complexity is easily controlled. Furthermore, no motion estimation is required which is known to be problematic when images are heavily corrupted by noise.

Our approach has been developed in order to denoise biological images sequences involved in the analysis of fluorescently tagged proteins moving around the Golgi apparatus and participating in the intra-cellular traffic. These proteins are embedded into vesicles whose movement is supposed to depend on a microtubules network.

2. METHOD DESCRIPTION

2.1. Model specification

We consider a statistical image model described by the following equation: $Y_i = u(\mathbf{x}_i) + \xi_i$ where $\mathbf{x}_i \in \Omega \subset \mathbb{R}^d$ with $d = 3$ for 2D image sequences or $d = 4$ for 3D

volume sequences. The image function $u_i = u(\mathbf{x}_i)$ has to be recovered from the data Y_i corrupted by an independent zero-mean random noise ξ_i with variance τ_i^2 .

We consider that u is a locally piecewise constant function in a neighborhood of a point ξ_i . However, the size and the shape of this neighborhood are unknown and must be estimated. Once characteristics of such a neighborhood of \mathbf{x}_i are determined, the regression function can be estimated by optimizing a local maximum likelihood (ML) criterion. The designed method addresses these issues and estimate u in an iterative way.

The main idea in our approach is to define a sequence of increasing spatio-temporal neighborhoods $(W_{i,n})_{n \in [0:N]}$ at each point \mathbf{x}_i such that $W_{i,n} \subset W_{i,n+1}$ and N indexes the larger window. The initialization of the window sequence is performed by setting an initial window $W_{i,0}$ at point \mathbf{x}_i . As an example, this first window may contain the 26 nearest neighbors in space domain. Then, we can compute an initial estimate of $u(\mathbf{x}_i)$, $\hat{u}_{i,0}$, and its associated variance $\hat{\sigma}_{i,0}^2$ as:

$$\hat{u}_{i,0} = \frac{1}{|W_{i,0}|} \sum_{\mathbf{x}_j \in W_{i,0}} Y_j \quad \text{and} \quad \hat{\sigma}_{i,0}^2 = \frac{1}{|W_{i,0}|^2} \sum_{\mathbf{x}_j \in W_{i,0}} \hat{\tau}_j^2 \quad (1)$$

where $\hat{\tau}_j^2$ is an empirical estimate of the local noise variance τ_j^2 (see Section 3). This initialization step provides the first estimates of the two sequences of estimates $(\hat{u}_{i,n})_{n \in [0:N]}$ and $(\hat{\sigma}_{i,n}^2)_{n \in [0:N]}$. The next step consists in considering a larger window $W_{i,1}$ with $W_{i,0} \subset W_{i,1}$. Then, at iteration $n = 1$, we calculate a new estimate $\hat{u}_{i,1}$ and $\hat{\sigma}_{i,1}^2$ over $W_{i,1}$. At iteration n , we consider weighted averages and get:

$$\hat{u}_{i,n} = \sum_{j \in W_{i,n}} \omega_{ij} Y_j \quad \text{and} \quad \hat{\sigma}_{i,n}^2 = \sum_{j \in W_{i,n}} \omega_{ij}^2 \hat{\tau}_j^2, \quad (2)$$

and the weights ω_{ij} are defined as a function of the contrast between the estimate $\hat{u}_{i,n-1}$ at point \mathbf{x}_i and the estimates $\hat{u}_{j,n-1}$ at points \mathbf{x}_j belonging to the neighborhood $W_{i,n}$.

2.2. Estimation step

Considering its robustness and smoothing properties, we have chosen the influence function of the Huber M-estimator defined as

$$f(x) = \begin{cases} 1 & \text{if } |x| \leq 1 \\ \frac{1}{|x|} & \text{else} \end{cases} \quad (3)$$

but other influence functions are possible [1]. At iteration n , the weights are calculated as follows:

$$\omega_{ij} = \frac{f(\delta_{ij})}{\sum_{\mathbf{x}_j \in W_{i,n}} f(\delta_{ij})} \quad (4)$$

where the normalized contrast δ_{ij} is defined by:

$$\delta_{ij} = \frac{\hat{u}_{i,n-1} - \hat{u}_{j,n-1}}{\lambda \hat{\sigma}_{i,n-1}}. \quad (5)$$

Therefore, the weights decide which points in the spatio-temporal neighborhood should contribute to the estimation. This decision is made under the hypothesis that the contrast is Gaussian distributed. The parameter λ controls the probability of false alarm to assign a point to the region $W_{i,n}$. In our experiments, we will take $\lambda = 3$ which corresponds to a probability of type I of 0.0027.

2.3. Window geometry

The second idea governing the definition of the sequence of neighborhoods $(W_{i,n})_{n \in [0:N]}$ is to combine the spatial neighborhood with the temporal neighborhood. This can be achieved not only by choosing a general shape/geometry for the neighborhoods but also by the way they are increased through iterations. Because the data are selected by weights, we can use a simple anisotropic hyper-cubic spatio-temporal volume for the geometry. This neighborhood can be easily parametrized respectively by its spatio-temporal extents. Note that, in confocal microscopy, $dz = 3dx = 3dy$ and this spatial anisotropy of 1/3 along the z dimension can be taken into account by scaling the depth dimension accordingly. As to the time dimension, we can prefer to separate the past and the future by introducing two temporal extents for each spatio-temporal neighborhood, *i.e.* t^- to denote the extent in the direction of the past and t^+ to denote the extent in the direction of the future. We can even apply a causal temporal filtering by setting $t^+ = 0$.

2.4. Local stopping rules

Two pointwise rules are used to stop the estimation process. The first one allows us to estimate the optimal window for the estimation of u_i . The criterion to measure the closeness of the estimator \hat{u} to the unknown function u is the local L_2 risk. It can be decomposed in two terms: squared bias and variance. The selected window will be the window which achieves an optimal compromise between these two terms. Letting $\mathbb{E}(\cdot)$ denote the mathematical expectation, the local L_2 risk can be written as:

$$\mathbb{E}[\hat{u}_{i,n} - u_i]^2 = [\text{Bias}(\hat{u}_{i,n})]^2 + \hat{\sigma}_{i,n}^2 \quad (6)$$

and the optimal solution corresponds to :

$$\mathbb{E}[\hat{u}_{i,n} - u_i]^2 \approx 2\hat{\sigma}_{i,n}^2. \quad (7)$$

Considering the sequence of growing windows $(W_{i,n})_{n \in [0:N]}$ and the sequence of estimates $(\hat{u}_{i,n})_{n \in [0:N]}$ of variance $(\hat{\sigma}_{i,n}^2)_{n \in [0:N]}$, the ideal window is the last one satisfying the following pointwise statistical rule [10, 11, 8]:

$$\forall n' < n, \quad |\hat{u}_{i,n} - \hat{u}_{i,n'}| < \eta \hat{\sigma}_{i,n'}. \quad (8)$$

This local statistical rule (8) means that, for each point \mathbf{x}_i , while the estimates $\hat{u}_{i,n}$ are sufficiently close to each other

then we decide to continue the estimation process. The factor η can be chosen in the range [2, 4] in order to adapt the decision rule [11, 3, 8].

We also introduce a second rule which consists in locally stopping the estimation if the variance is small enough. This upper bound is set considering that we are dealing with digital integer data. Typically, if $\sigma_{i,n} < 1/3$, we stop the estimation for the point \mathbf{x}_i . This simple rule enables to save computing time for uniform regions where larger windows will be accepted with a high probability.

3. IMPLEMENTATION

As noticed in subsection 2.1, an estimation of the noise variance is required. For denoising purposes, we assume that the noise variance is constant over the whole volume sequence *i.e.* $\tau_i^2 = \tau^2$ and that it can be robustly estimated by calculating pseudo-residuals ε_i defined at point \mathbf{x}_i in a similar way than in [13] by : $\varepsilon_i = 1/\sqrt{72} (8Y_i - \Delta Y_i)$ where ΔY is the discrete Laplacian of Y calculated over a $4D$ 8-neighborhood. Given the residuals ε_i , we can then robustly estimate the noise variance τ by:

$$\tau = 1.4826 \text{ Med}_i (| \varepsilon_i - \text{Med}_j | \varepsilon_j |). \quad (9)$$

A pointwise estimation of the variance is also proposed in [8]. We have also seen in sub-sections 2.2 and 2.3 that λ and η can be well calibrated using statistical arguments and set to 3 and $2\sqrt{2}$ in experiments. The maximum number of iterations N is set to 8 which is a good compromise between computation cost and accuracy.

4. EXPERIMENTS

Our method has been tested on both synthetic and experimental $2D$ and $3D$ image sequences using fast $3D$ microscopy deconvolution[14]. The method is applied to fluorescently tagged proteins moving around the Golgi apparatus and participating to the intra-cellular traffic. Figure 1 contains the temporal intensity profiles extracted for three labeled pixels. They correspond to three typical classes of profiles observed in those sequences. Figure 3 shows slices of the $2D + t$ volume, that facilitate a qualitative evaluation of the filtering process. Thus, we can see that discontinuities are well preserved and in the same time, that the sequence is strongly smoothed. In addition, simulations in table 1 demonstrate that $PSNR = 20 \log_{10}(255/rms)$ where rms denotes the quadratic error, is drastically improved. Finally, figures 4 (a) et (b) show the interest of a denoising processing for visualization purposes.

τ	$PSNR_{noisy}$	$PSNR_{denoised}$	$IPSNR$	rms
15	24.41	35.07	10.65	4.45
20	22.00	33.88	11.88	5.12

Table 1. Filtering performance for two noise levels. The test sequence was obtained adding a white Gaussian noise of variance τ^2 to pre-filtered data in order deal with realistic dynamics and photometry. $N = 8$, $\lambda = 3$, $\eta = 2\sqrt{2}$.

5. CONCLUSION

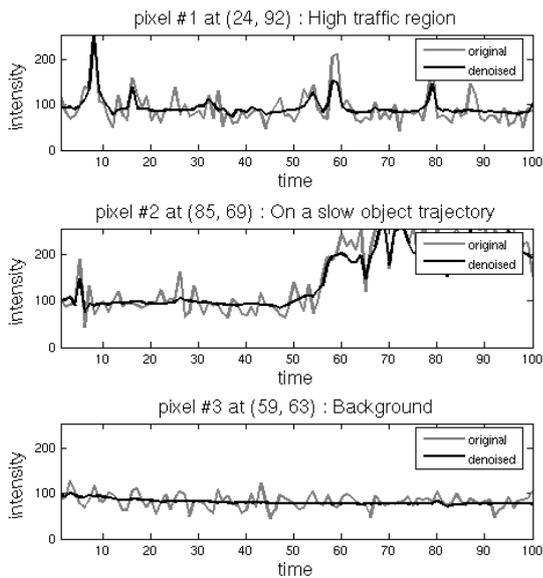
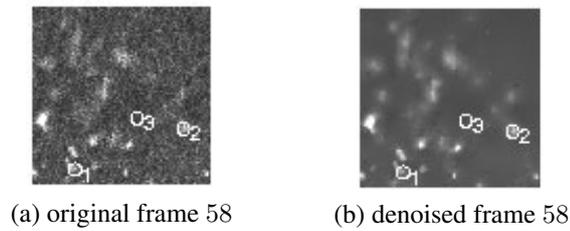
We have presented a original adaptive and efficient method for denoising $2D + t$ or $3D + t$ image sequences. Numerous experiments demonstrate its ability to smooth images while preserving discontinuities in time as well as in space. Extensions of this method is to reduce the computational load, and to exploit spatio-temporal features (window size) given by the algorithm. Such a denoising algorithm is a first step that will allow us to extract information about the lifetime kinetics of specific Rab proteins. This novel approach does not require motion compensation and can be further used for biological studies where dynamics have to be analyzed.

6. ACKNOWLEDGMENTS

The authors would like to thank the ACI-IMPBio (MODYNCCELL5D project) for its support and Curie Institute for providing image sequences.

7. REFERENCES

- [1] M.J. Black, G. Sapiro, D.H. Marimont, and D. Heeger, "Robust anisotropic diffusion", *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 421–432, 1998.
- [2] J.C. Brailean, R.P. Kleihorst, S. Efstratiadis, A.K. Katsaggelos, and R.L. Lagendijk, "Noise reduction filters for dynamic image sequences: a review," *Proc. of the IEEE*, vol. 83, no. 9, pp. 1272–1291, 1995.
- [3] V. Katkovnik, K. Egiazarian, and J. Astola, "Adaptive window size image denoising based on intersection of confidence intervals (ICI) rule". *J. Math. Imag. Vis.*, vol. 16, no. 3, pp. 223–235, 2002.
- [4] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, no. 1-4, pp. 259–268, 1992.
- [5] Suk Ho Lee and Moon Gi Kang, "Spatio-temporal video filtering algorithm based on 3-D anisotropic diffusion equation," in *Proc. of IEEE Int. Conf. Image Process. (ICIP'98)*, Chicago, 1998, vol. 3(2), pp. 447–450.
- [6] D. Uttenweiler, C. Weber, B. Jähne, R.H. Fink, and H. Schar, "Spatiotemporal anisotropic diffusion filtering to improve signal to noise ratios and object restoration in fluorescence microscopic image sequences," *J. Biomedical Optics*, vol. 8, no. 1, pp. 40–47, 2003.
- [7] N. Rajpoot, Z. Yao, and R. Wilson, "Adaptive wavelet restoration of noisy video sequences," in *Proc. of IEEE Int. Conf. Image Process. (ICIP'04)*, Singapore, 2004.



(c) temporal signals associated to the 3 marked pixels

Fig. 1. Three pixels have been selected in frame number 58 extracted from a $2D + t$ volume.

- [8] C. Kervrann, "An adaptive window approach for image smoothing and structures preserving," in *Proc. of Eur. Conf. Comp. Vis. (ECCV'04)*, Prague, Czech Republic, 2004, pp. 132–144.
- [9] Danny Barash, "A fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 24, no. 6, pp. 844–847, 2002.
- [10] O. Lepski, "Asymptotically minimax adaptive estimation 1: upper bounds," *SIAM J. Th. Probab. and Appl.*, vol. 36, no. 4, pp. 654–659, 1991.
- [11] J. Polzehl and V.G. Spokoiny, "Adaptive weights smoothing with applications to image restoration," *J. Royal Statist. Soc.: Series B (Stat. Meth.)*, vol. 62, no. 2, pp. 335–354, 2000.
- [12] L. Stankovic, "Performance analysis of the adaptive algorithm for bias-to-variance tradeoff," *IEEE Trans. on Signal Processing*, vol. 52, no. 5, pp. 1228–1234, 2004.
- [13] T. Gasser, L. Sroka, and C. Jennen Steinmetz, "Residual variance and residual pattern in nonlinear regression," *Biometrika*, pp. 625–633, 1986.
- [14] J.-B. Sibarita, H. Magnin, and J.R. De Mey, "Ultra-fast 4D microscopy and high throughput distributed deconvolution," in *Proc. of IEEE Int. Symp. on Biomedical Imaging (ISBI'02): Macro to Nano*, Washington, 2002, pp. 769–772.

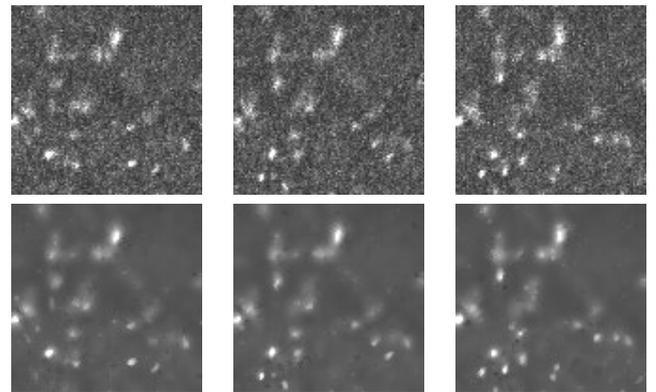


Fig. 2. Three frames of a real $2D$ image sequence denoised with our method.

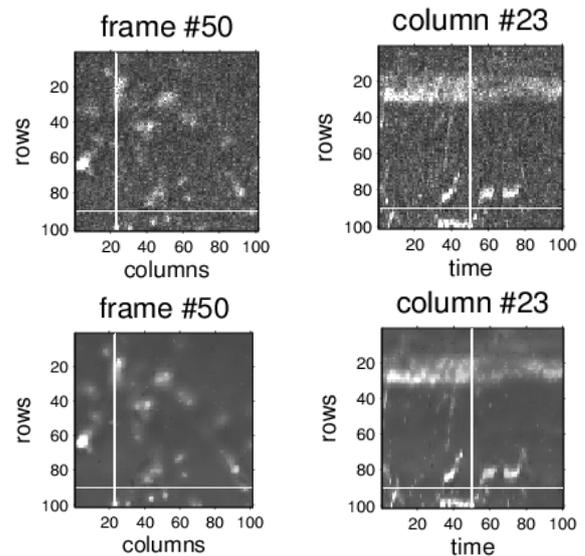


Fig. 3. Noisy and denoised data (see figure 2) viewed as a $3D = 2D + t$ sliced volume. Temporal discontinuities are well preserved on slice $column = 23$.

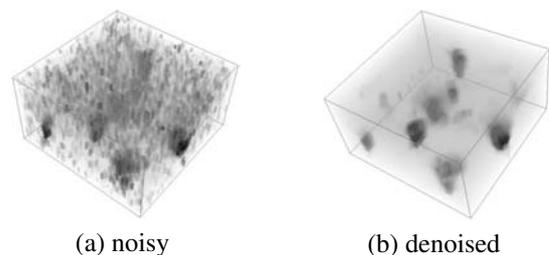


Fig. 4. Volume rendering of a cropped region of a $3D$ frame extracted from a $4D$ volume. Figure (a) shows raw data while figure (b) shows the restoration results using our $4D$ spatio-temporal filtering method.