

# Adaptive Spatio-Temporal Restoration for 4D Fluorescence Microscopic Imaging

Jérôme Boulanger<sup>1,2</sup>, Charles Kervrann<sup>1,2</sup>, and Patrick Bouthemy<sup>1</sup>

<sup>1</sup> IRISA - INRIA, Campus Universitaire de Beaulieu, 35042 Rennes, France

<sup>2</sup> INRA - MIA Domaine de Vilvert 78352 Jouy en Josas, France

**Abstract.** We present a spatio-temporal filtering method for significantly increasing the signal-to-noise ratio (SNR) in noisy fluorescence microscopic image sequences where small particles have to be tracked from frame to frame. Image sequence restoration is achieved using a statistical approach involving an appropriate on-line window geometry specification. We have applied this method to noisy synthetic and real microscopic image sequences where a large number of small fluorescently labeled 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 exploited for biological studies where the dynamics of small objects of interest have to be analyzed in molecular and sub-cellular bio-imaging.

## 1 Introduction

Time-lapse confocal microscopy is now a popular technology in biomedicine and biophysics for analysis of dynamics at sub-cellular or molecular levels. Several applications in biology are concerned with the acquisition of sequences of  $2D$  images or  $3D$  volumes representing small fluorescently-tagged particles with highly variable velocities. In order to analyze these moving objects, the acquisition time is reduced and consequently the signal-to-noise ratio becomes low. As a consequence, motion estimation and trajectography of such small objects remain difficult tasks.

Several algorithms have been developed for  $3D$  image sequences restoration. Traditionally, these  $4D$  data sets are processed, volume by volume, using  $3D$  spatial filters, in order to save computation time, even if spatio-temporal filtering would appear more appropriate. In contrast, spatio-temporal filtering has been largely investigated in video processing, and usually relies on a motion-compensated approach [1]. Most of these spatio-temporal filters can nevertheless be considered as an extension of well-known principles developed for  $2D$  image restoration or smoothing. To our knowledge, non-linear anisotropic diffusion [2, 3], recent PDE-based methods [4], Wiener filters [5], and wavelet shrinkage [6], have been used for processing  $2D$  image sequences. Finally,  $3D$  image sequence smoothing algorithms have been specifically developed for ultrasound imaging [7, 8].

In this paper we address the problem of  $3D$  image sequence restoration by significantly extending the framework described for smoothing  $2D$  still images

in [9, 10, 11]. Unlike robust anisotropic diffusion [12] and non-linear Gaussian filtering [13], this local adaptive estimation approach provides a scale selection for each pixel by locally estimating the appropriate spatial filtering window [9]. Additionally, a confidence level (i.e. variance) attached to the each pixel is provided. Furthermore, the *Total Variation* (TV) minimization method [14], commonly-used for 2D image restoration, cannot be easily extended to a space-time domain.

The novel method introduced in this paper is able to jointly estimate the spatial and temporal discontinuities in 3D image sequences. Also, the proposed algorithm is non-iterative and relatively fast since its complexity can be bounded by the user, which is necessary for processing large 4D data sets. Finally, the theoretical properties of such an estimator are well known in non-parametric statistics for adaptive smoothing [15, 11, 10].

While being general, the described method has been designed for analyzing<sup>1</sup> the role of fluorescence-tagged proteins moving around the Golgi apparatus and participating to the intra-cellular traffic. These proteins are embedded into vesicles whose movement is supposed to be dependent on a microtubules network. These vesicles propelled by motor proteins follow these polarized “cables”. This mechanism explains the observed high velocities which could not be accounted for a simple diffusion. Let us point out that the method is data-driven and requires no motion estimation, which is known to be problematic when images are heavily corrupted by noise.

## 2 Proposed Approach

We consider the following statistical image model:  $Y_i = u(\mathbf{x}_i) + \xi_i$ , where  $\mathbf{x}_i \in \Omega$  represents the pixel location in the image domain  $\Omega \subset \mathbb{R}^d$  ( $d = 3$  for 2D and  $d = 4$  for 3D image sequences). The image function  $u_i = u(\mathbf{x}_i)$  has to be recovered from observations  $Y_i$ . The errors  $\xi_i$  are assumed to be independent zero-mean Gaussian variables with unknown variances  $\tau_i^2$ .

We need minimal prior assumptions on the structure of the image for recovering  $u$ . In what follows, we assume that  $u(\mathbf{x}_i)$  is a locally piecewise constant function in the neighborhood of the pixel  $\mathbf{x}_i$ . However, the size and shape of these neighborhoods is not constant over the image sequence and must be estimated too. Once such a neighborhood has been determined, the regression function  $u$  can be estimated by optimizing a local maximum likelihood (ML) criterion. The proposed method addresses these two issues as described below.

One important feature of our approach is to define a sequence of increasing nested spatio-temporal neighborhoods  $(W_{i,n})_{n \in [0:N]}$  at each point  $\mathbf{x}_i$ , i.e.  $W_{i,n} \subset W_{i,n+1}$  with  $N$  indicating the largest window. At the initialization, we choose the 26 nearest neighbors in the 3D space domain as the *pilot* (starting) window  $W_{i,0}$  at point  $\mathbf{x}_i$ ,  $\forall \mathbf{x}_i \in \Omega$ . Then, we can compute an initial estimate  $\hat{u}_{i,0}$  of  $u(\mathbf{x}_i)$  and its associated variance  $\hat{\sigma}_{i,0}^2$  as:

---

<sup>1</sup> The authors would like to thank the ACI-IMPBio for its support and the Curie Institute for the image sequence data-base.

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

where  $|W_{i,0}|$  denotes the number of pixels in  $W_{i,0}$  and  $\hat{\tau}_j^2$  is an empirical estimate of the local noise variance  $\tau_j^2$  as described in Section 3. This initialization step provides the first estimates of the two sequences  $(\hat{u}_{i,n})_{n \in [0:N]}$  and  $(\hat{\sigma}_{i,n}^2)_{n \in [0:N]}$ . The next step consists in considering a larger window  $W_{i,1}$  such that  $W_{i,0} \subset W_{i,1}$  and calculating new estimates  $\hat{u}_{i,1}$  and  $\hat{\sigma}_{i,1}^2$  over  $W_{i,1}$ . At iteration  $n$  we define the estimator as:

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

where the weights  $\omega_{ij}$  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  $W_{i,n}$ . The weights  $\omega_{ij}$  and the geometry and size of the windows  $W_{i,n}$  will be formally defined in the next Sections.

## 2.1 Adaptive Weights

At iteration  $n \geq 1$ , the weights are calculated as follows:

$$\omega_{ij} = \frac{f(\delta_{ij})}{\sum_{\mathbf{x}_j \in W_{i,n}} f(\delta_{ij})} \quad \text{and} \quad \delta_{ij} = \frac{\hat{u}_{i,n-1} - \hat{u}_{j,n-1}}{\lambda \hat{\sigma}_{i,n-1}} \quad (3)$$

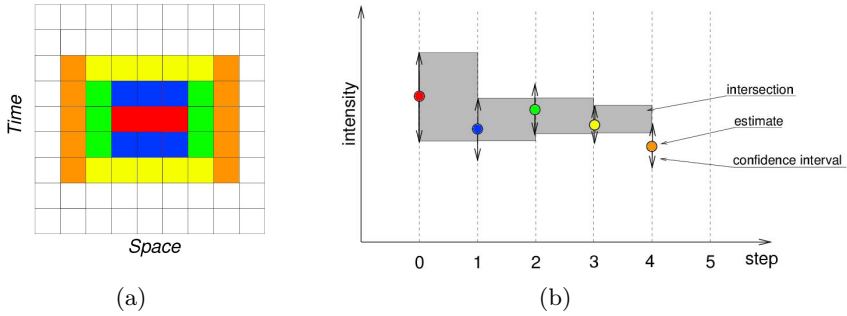
where  $f$  is a function of the normalized contrast  $\delta_{ij}$ . Considering its robustness and smoothing properties, we have chosen the influence function of the Huber M-estimator defined as:

$$f(x) = \begin{cases} 1 & \text{if } |x| \leq 1 \\ \frac{1}{|x|} & \text{otherwise} \end{cases} \quad (4)$$

but other influence functions are possible [12]. Therefore, the weights decide which points  $\mathbf{x}_j$  in the spatio-temporal neighborhood  $W_{i,n}$  should contribute to the estimation of  $u(\mathbf{x}_i)$ . This decision is made under the hypothesis that the contrast is Gaussian distributed. Hence, the parameter  $\lambda$  controls the probability of false alarm (satisfy the test when it should not) to include a point  $\mathbf{x}_j$  in  $W_{i,n}$ . In our experiments, we set  $\lambda = 3$  which corresponds to a commonly-used probability of error of type I of 0.036.

## 2.2 On-Line Window Geometry Specification

One of the main contributions of this work is the on-line adaptation of the neighborhood sequence  $(W_{i,n})_{n \in [0:N]}$ . It is worth noting that points  $\mathbf{x}_j \in W_{i,n}$  that contribute to the estimation of  $u(\mathbf{x}_i)$  are thus selected by weights in a quite flexible and efficient way unlike usual methods which exploit geometry-based design of the neighborhood. This allows us to use a simple hyper-cubic spatio-temporal



**Fig. 1.** (a) Spatio-temporal neighborhood: colors correspond to iterations plotted in (b); (b) confidence intervals: circles represent estimates  $\hat{u}_{i,n}$  obtained at each iteration  $n$ . The grey rectangles represent the intersection between the current confidence interval and the previous one. As long as the estimate belongs to this intersection, the estimation process is updated.

volume for the window shape. Nevertheless, we separate the space dimension from the time dimension and parametrize the spatio-temporal volume by introducing two extents. Figure 1 shows the considered neighborhood sequence and how neighborhoods are enlarged along with the iterations. It is worth stressing that this figure only represents a particular case and that the sequence of neighborhoods is not defined *a priori* but locally estimated simultaneously to the image filtering procedure. Thus, each point is assigned an adapted smoothing support. This support is estimated according to the stopping rules described in the next section.

### 2.3 Window Estimation

Two pointwise rules are used to drive the window estimation process. The first one is designed to estimate the optimal window at  $\mathbf{x}_i$  and is based on the measure of the closeness of the estimator  $\hat{u}$  to the unknown function  $u$  given by the local  $L_2$  risk. This measure of performance can be decomposed in two terms, that is the squared bias and the variance as  $\mathbb{E}[\hat{u}_{i,n} - u_i]^2 = [\text{bias}(\hat{u}_{i,n})]^2 + \hat{\sigma}_{i,n}^2$ , where  $\mathbb{E}(\cdot)$  denotes the mathematical expectation. In what follows, we reasonably assume the squared bias is an increasing function of the window size and the variance is a decreasing function of the window size [9, 15, 11, 10]. Then, the selected window will be the window which achieves an optimal compromise between these two terms, i.e. the squared bias and variance terms must be of the same order:  $\text{bias}(\hat{u}_{i,n}) \asymp \hat{\sigma}_{i,n}$ . A closed-form optimal solution is not available for our non-linear estimator and we arbitrarily define, as usually done, the optimal window as the window for which the squared bias and the variance are nearly the same. Now let us consider the sequence of increasing windows  $(W_{i,n})_{n \in [0:N]}$  and the sequence of estimates  $(\hat{u}_{i,n})_{n \in [0:N]}$  of variance  $(\hat{\sigma}_{i,n}^2)_{n \in [0:N]}$ . A practical rule based on pairwise comparison of successive estimates can be derived to detect the optimal window. We define the largest window satisfying the following pointwise statistical rule [9, 15, 11]:

$$|\hat{u}_{i,n} - \hat{u}_{i,n'}| < \eta \hat{\sigma}_{i,n'}, \quad \forall n' < n, \quad (5)$$

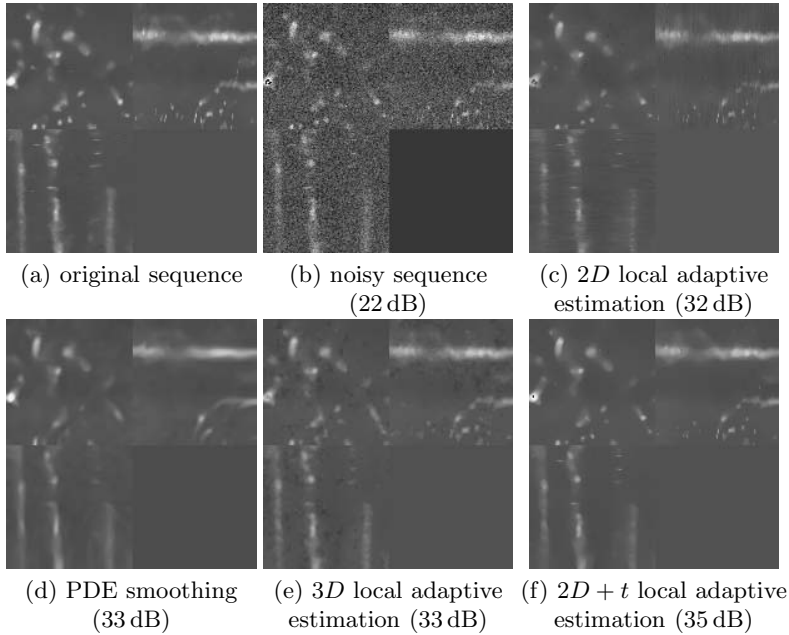
as the optimal window since  $W_{i,n'} \subset W_{i,n}$  at  $\mathbf{x}_i$ . Strictly speaking, while the estimates  $(\hat{u}_{i,n'})_{n' \leq n}$  are sufficiently close to each other, then we decide to continue the estimation process. The factor  $\eta$  can be easily chosen in the range  $[2, 4]$  in order to adapt the decision rule as justified in [9, 11, 10]. As shown in Figure 1, this rule amounts to continue the pointwise estimation process, while new estimates belong to the intersection of estimated confidence intervals  $[\hat{u}_{i,n} - \eta \hat{\sigma}_{i,n}, \hat{u}_{i,n} + \eta \hat{\sigma}_{i,n}]$ . This representation brings to the fore that we don't have to store all the previous estimates  $(\hat{u}_{i,n'})_{n' \leq n}$  but only the intersection of confidence intervals, the previous estimate and its variance for each point. We also introduce a second rule which consists in locally stopping the pointwise estimation if the confidence intervals is small enough: when the output  $\hat{u}_{i,N}$  is stored in an integer format, it is useless to continue to improve the estimation when the confidence associated to the estimate is below the unit. This second rule enables to autonomously stop the estimation process and makes the algorithm more data-driven.

### 3 Algorithm Implementation

As noticed in Section 2, an estimation of the noise variance is required. Here, we assume that the noise variance is constant over the whole 3D image sequence. It can be robustly estimated by calculating pseudo-residuals  $\varepsilon_i$  as described in [16]. If we choose a 4D 8-neighborhood, pseudo-residuals are compactly represented by  $\varepsilon_i = (8Y_i - \Delta Y_i) / \sqrt{72}$  where  $\Delta Y_i$  is the discrete Laplacian at  $\mathbf{x}_i$  and the constant  $\sqrt{72}$  is introduced to insure that  $\mathbb{E}[\varepsilon_i^2] = \tau^2$ . Given the residuals  $\varepsilon_i$ , we can then robustly estimate the noise variance  $\tau^2$  by:  $\tau = 1.4826 \text{ med}_i (|\varepsilon_i - \text{med}_j |\varepsilon_j| |)$ . A local estimation of the noise variance is proposed in [9] and can also be used when the noise model proposed in Section 2 is not appropriate to describe the spatio-temporal inhomogeneity of the image sequence. Moreover, for Poisson distributed noises, the Anscombe transform can be applied [17]. We have also seen in Sections 2.1 and 2.2 that  $\lambda$  and  $\eta$  can be well calibrated using statistical arguments. In our experiments,  $\lambda$  is set to 3 and  $\eta$  to  $2\sqrt{2}$ . During the estimation, we alternate the increasing of the spatial and temporal extents of the 4D spatio-temporal neighborhoods. Furthermore, the algorithm can be easily parallelized. Actually, estimation steps use only local information and have been distributed over several CPUs. Finally, another possibility to speed up the algorithm is to use a dyadic scheme when increasing the extent of the neighborhood.

### 4 Experiments

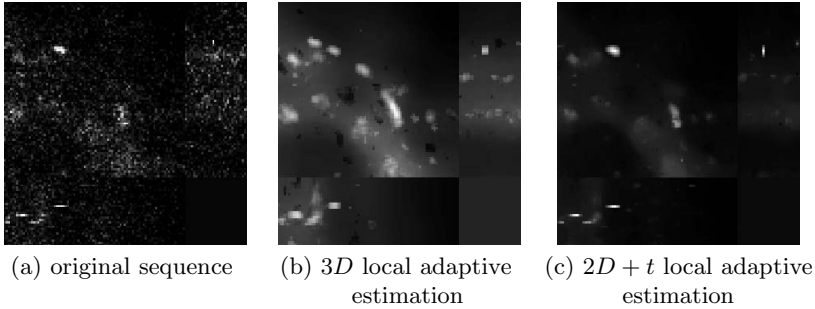
The proposed method has been applied to both synthetic and experimental sequences of 2D and 3D images obtained by fast 3D deconvolution microscopy [18]. Since the dynamics of vesicles are unknown, we first simulate a noisy sequence of 2D images by adding a realistic white Gaussian noise to a real denoised



**Fig. 2.** Comparison with other method on a synthetic sequence. Six  $2D$  image sequences are represented as  $100 \times 100 \times 100$  volumes. First three quadrants correspond slice  $XY$ ,  $TY$  and  $XT$ . The last one is left empty. (a) original sequence used as the ground truth ; (b) artificially noisy sequence ; (c) spatial adaptive estimation [9], i.e. each image of the sequence is processed independently ; (d)  $3D$  anisotropic diffusion [4] when the sequence is considered as a homogeneous  $3D$  volume ; (e) adaptive estimation algorithm applied the sequence considered as a  $3D$  volume ; (f) our  $2D + t$  adaptive estimation (see text).

image sequence (Figs 2(a) – 2(b)). We will also compare our method to other methods on this simulated sequence.

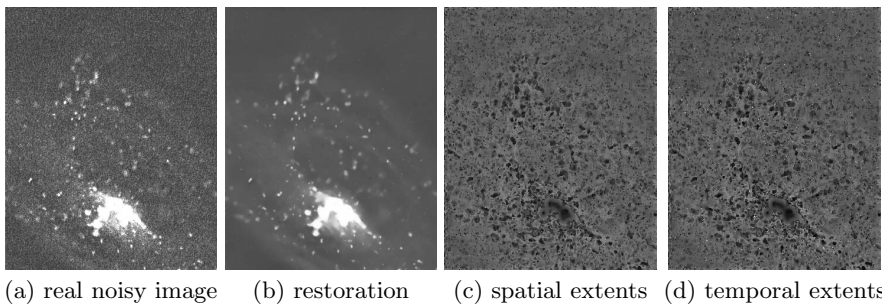
In Fig. 2(c), some blinking effects or strides are visible on spatio-temporal slices when each  $2D$  frame of the sequence are processed independently with the adaptive estimation described in [9]. Anisotropic diffusion [4] has been applied on the  $2D$  sequence considered as a  $3D$  volume. In Fig.2(d), the sequence is strongly smoothed and some details are lost. In Fig.2(e), the  $3D$  adaptive estimation algorithm described in [9] has been applied to the  $3D$  volume. As expected, considering the temporal dimension as a spatial dimension is not appropriate to cope with temporal discontinuities. Finally, the results of our method are shown in Fig.2(f). The peak-signal-to-noise-ratio defined in decibels ( $dB$ ) as  $PSNR = 20 \log_{10}(255/mse)$  where  $mse$  denotes the mean squared error between the original sequence and the result of filtering process, is drastically improved and the dynamics of particles are well preserved. Figure 3(a) shows a detail of a real  $2D$  image sequence. It is first processed by equally considering the three space-time dimensions. Undesirable blurring over three successive images of a



**Fig. 3.** Detail of a sequence denoised with two versions of the filter:  $3D$  estimator and the proposed  $2D + t$  estimator. First three quadrants correspond slice  $XY$ ,  $TY$  and  $XT$ . The last one is left empty.

single spot along its track are visible (Fig. 3(b)). In the case of spatio-temporal filtering (Section 2), these artifacts are removed as shown in Fig. 3(c).

We have also applied our adaptive spatio-temporal filtering method on a database of video-microscopy sequences in order to facilitate the automatic analysis of the dynamics of two proteins Rab6a and Rab6a' involved in particular steps of membrane trafficking. These proteins are fluorescently-tagged and move around the Golgi apparatus. The function of this organelle is to package materials for export from the cell. Figures 4(a) and (b) show respectively a real noisy image extracted from a video-microscopy sequence ( $360 \times 445 \times 10$  pixels) and the corresponding restored image for the fifth  $2D$  slice in depth. The algorithm provides visually satisfying results after ten iterations as demonstrated in Figs. 5 (a) and (b). that illustrate a cropped region of a  $3D$  image. Note that, the point-wise window extents shown in Figs. 4 (c) and (d) roughly correspond to the size



**Fig. 4.** One frame depicting the spot positions at time  $t = 2s$ . The dynamics of fluorescent tags has been recorded by fast  $3D$  deconvolution microscopy. The large white region corresponds to the Golgi apparatus while small spots are vesicles moving with a high average speed ( $\sim 10$  pixels/frame). (a) real noisy image ( $2D$  slice of  $3D$  frame); (b) denoised image using the spatio-temporal adaptive estimation ; (c) spatial extents of windows ; (d) time extents of windows (small extents are coded in black and are located in the vicinity of static or moving spots).



**Fig. 5.** Volume rendering of a cropped region of a 3D image extracted from the sequence in Fig 4. (a) original real noisy 3D frame; (b) results using our adaptive spatio-temporal method.

of space-time structures in image sequences. These spatio-temporal features will be further exploited for analysis of these biological data. The processing time of a sequence of 50 2D frames ( $360 \times 445$ ) is typically 10 min

## 5 Conclusion

We have described an original, adaptive and efficient method for 2D and 3D image sequence restoration. Experiments on synthetic and real video-microscopy image sequences have demonstrated its ability to smooth images while preserving both spatial and temporal discontinuities. Moreover, the presented method does not require a fine tuning of the parameters which are well calibrated using statistical arguments. This approach does not require motion compensation and can be further used for biological studies where dynamics have to be analyzed in noisy sequences. In particular, we could exploit spatio-temporal features (window size) provided by the algorithm, which can be regarded as a preliminary step for analysis of the lifetime kinetics of specific Rab proteins. Furthermore, this method is not restricted to biological applications and could be applied for smoothing three dimensional ultrasound images using an estimator adapted to the Rayleigh noise model.

## References

1. Brailean, J., Kleihorst, R., Efstratiadis, S., Katsaggelos, A., Lagendijk, R.: Noise reduction filters for dynamic image sequences: A review. *Proc. of the IEEE* **83** (1995) 1272–1291
2. Lee, S.H., Kang, M.G.: Spatio-temporal video filtering algorithm based on 3-D anisotropic diffusion equation. In: *ICIP'1998*. Volume 3(2)., Chicago, (1998) 447–450
3. Uttenweiler, D., Weber, C., Jähne, B., Fink, R., Scharr, H.: Spatiotemporal anisotropic diffusion filtering to improve signal to noise ratios and object restoration in fluorescence microscopic image sequences. *Jour. of Biomedical Optics* **8** (2003) 40–47
4. Tschumperlé, D., Deriche, R.: Vector-valued image regularization with pde's : A common framework for different applications. *IEEE Trans. on Pattern Analysis and Machine Intelligence* **27** (2005) 506–517



5. Dekeyser, F., Bouthemy, P., Pérez, P.: Spatio-temporal wiener filtering of image sequences using a parametric motion model. In: ICIP'2000, Vancouver, Canada (2000)
6. Rajpoot, N., Yao, Z., Wilson, R.: Adaptive wavelet restoration of noisy video sequences. In: ICIP'2004, Singapore (2004)
7. Angelini, E., Laine, A., Donis, J., Gersony, D., Homma, S.: Quantification of right and left ventricular function with real time three-dimensional ultrasound. In: Proc. of the 23rd Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society. (2001)
8. Mikula, M., Preusser, T., Rumpf, M.: Morphological image sequence processing. *Computing and Visualization in Science* **6** (2004) 197–209
9. Kervrann, C.: An adaptive window approach for image smoothing and structures preserving. In: Proc. of 8th Eur. Conf. on Computer Vision, ECCV'2004, Prague, Czech Republic (2004) 132–144
10. Katkovnik, V., Egiazarian, K., Astola, J.: Adaptive window size image denoising based on intersection of confidence intervals (ici) rule. *Jour. of Mathematical Imaging and Vision* **16** (2002) 223–235
11. Polzehl, J., Spokoiny, V.: Adaptive weights smoothing with applications to image restoration. *Jour. of the Roy. Stat. Soc.: Series B (Stat. Meth.)* **62** (2000) 335–354
12. Black, M., Sapiro, G., Marimont, D., Heeger, D.: Robust anisotropic diffusion. *IEEE Trans. on Image Processing* **7** (1998) 421–432
13. Barash, D.: A fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation. *IEEE Trans. on PAMI* **24** (2002) 844–847
14. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Phys. D* **60** (1992) 259–268
15. Lepski, O.: Asymptotically minimax adaptive estimation 1: upper bounds. *SIAM Jour. Theory of Probability and Application* **36** (1991) 654–659
16. Gasser, T., Sroka, L., Jennen Steinmetz, C.: Residual variance and residual pattern in nonlinear regression. *Biometrika* (1986) 625–633
17. Starck, J., Murtagh, F., Bijaoui, A.: *Image Processing and Data Analysis, the Multiscale Approach*. Cambridge University Press (2000)
18. Sibarita, J.B., Magnin, H., De Mey, J.: Ultra-fast 4D microscopy and high throughput distributed deconvolution. In: Proc. of IEEE Int. Symp. on Biomedical Imaging, ISBI'2002: Macro to Nano, Washington, (2002) 769–772