



HAL
open science

Simulation and estimation of fluorescence microscopy image sequences for intracellular dynamics and trafficking

Jérôme Boulanger, Charles Kervrann, Patrick Bouthemy

► **To cite this version:**

Jérôme Boulanger, Charles Kervrann, Patrick Bouthemy. Simulation and estimation of fluorescence microscopy image sequences for intracellular dynamics and trafficking. MICCAI'06, Oct 2006, Copenhagen, Denmark. hal-02753571

HAL Id: hal-02753571

<https://hal.inrae.fr/hal-02753571>

Submitted on 3 Jun 2020

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.

Simulation and Estimation of Fluorescence Microscopy Image Sequences for Intracellular Dynamics and Trafficking

Jérôme Boulanger^{1,2}, Charles Kervrann^{1,2}, Patrick Bouthemy¹

¹ IRISA - INRIA, Campus Universitaire de Beaulieu, 35042 Rennes – France

² INRA - MIA, Domaine de Vilvert, 78352 Jouy en Josas – France

Abstract— Image sequence analysis in video-microscopy for life sciences now has gained importance since molecular biology is presently having a profound impact on the way research is being conducted in medicine. However, image processing techniques that are currently used for modeling intracellular dynamics are still relatively crude. Indeed, complex interactions between a large number of small moving particles in a complex scene cannot be easily modeled, which limits the performance of object detection and tracking algorithms. This motivates our present research effort which is to develop a general estimation/simulation framework able to produce image sequences showing small moving spots in interaction and with variable velocities, corresponding to intracellular dynamics and trafficking in biology. It is now well established that spot trajectories can play a role in analysis of living cell dynamics and simulate realistic image sequences is then of major importance. We demonstrate the potential of the proposed simulation/estimation framework in experiments, and show that this approach can be also used to evaluate the performance of object detection/tracking algorithms in video-microscopy.

I. INTRODUCTION

A. Context in biology

The development of *system biology* is characterized by the settlement of new techniques and technologies producing a vast amount of data of different types or origins. Only automatic approaches for analysis and interpretation of complex and massive data will allow researchers to face this new challenge. This is already well established for a number of biological fields such as DNA sequence analysis, expression data analysis, DNA micro-arrays analysis, ... Also, in dynamical imaging of biological samples substantial amount of work is necessary to overcome conceptual and technological obstacles. This motivates our present research effort which is to develop novel techniques based on recent techniques in computer vision and signal processing to analyze information from 4D data related to intracellular dynamics and trafficking.

In video-microscopy, methods that estimate trajectories of small objects (chromosomes, vesicles, ...) of interest may encounter difficulties if the number of objects is large and the signal-to-noise ratio is low. Moreover, the tracked objects are not always visible in the sequence when tagging molecules separate suddenly from the target objects. Yet, most of the time the complexity of the dynamic processes involving many objects or groups of objects interacting, cannot be easily modeled. Finally, the corpus of data to be considered for a

comparative analysis in an experiment, formed by multiple image series acquisitions, is also massive. Nevertheless, it is now clear that the localization and spatio-temporal conformation of a large number of molecular constructions within the cell, their dynamic response to diverse chemical, physical or bio-molecular perturbations, are key elements for understanding the basic function mechanisms in life sciences. Typically, motion information and trajectories have to be extracted in order to analysis the dynamic response of the cell to the different perturbations and experimental conditions.

Then, we propose a simulation/estimation framework going in this research direction, able to model complex data corresponding to interactions between moving particles/spots with variable velocities. The idea is to propose parsimonious models of fluorescence microscopy image sequences, able to summarize complex data into a low dimensional set of parameters. These models will be exploited to generate artificial image sequences that mimic dynamics observed in real sequences.

B. Needs for simulation tools

In many application fields such as medical imaging or astronomy, simulations are required for validating physical models and understanding recorded data. In this section, we give the rational behind the idea of simulation methods for video-microscopy.

First, realistic simulations of dynamical processes usually give a qualitative representation of the observed spatio-temporal biological event. Simulation can be then considered as a computational tool that can help to understand some mechanisms of internal components within the cell. By interacting with the control parameters, an expert can artificially simulate processes close to the reality provided the dynamical models are known; this philosophy has been successfully exploited to understand dynamics of microtubule networks [1], [2]. By minimizing the difference between a set of descriptors computed from a real image sequence and the same set of descriptors computed from a simulated sequence, the parameters of the simulation method can be tuned to obtain an artificial sequence that reveals the same dynamical characteristics than the observed sequence. This set of estimated control parameters can be then considered a parsimonious representation of the underlying process.

Moreover, dynamical information extraction usually relies on tasks such as object detection, optical flow estimation or object tracking [3], [4]. These tasks cannot be done manually,

and they must be fast, reliable and reproducible. Furthermore, comparing object tracking results to a ground truth is the more straightforward method to assess performance of the applied method. Accordingly, simulation of a reliable ground truth is an important and challenging task especially in biomedical imaging. We point out that benchmarking data sets are for instance widely used to compare methods for image restoration [5] and optical flow estimation [6]. In video-microscopy, the photometric and dynamic characteristics of benchmarking data sets actually are not able to describe complex interactions between objects observed in real image sequences; in the case of tracking vesicles within the living cell in video-microscopy image sequences, *random walks* combined with parametric models are used for validation [3], which does not completely describe the movements of real moving objects in videos.

C. Our approach

In this paper, we then propose a powerful benchmarking method for simulating complex video-microscopy image sequences. We propose a realistic image sequence modeling framework describing the dynamical and photometric contents of video-microscopy image sequences. Unlike the biophysical approach which aims at describing the underlying physical phenomena [1], [2], the proposed approach is only based on the analysis of original image sequences. While being quite general, the described method has been designed for analyzing 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 microtubule network. These vesicles propelled by motor proteins move along these polarized “cables”. This mechanism explains the observed high velocities which could not be accounted by basic diffusions.

The remainder of this paper is organized as follows. In Section II, we give the properties that a simulation framework should satisfy to be performant. We adopt a two-step scheme: in Section III, a dynamical background model is described and a method is proposed to estimate the model parameters; in Section IV, a photometric and dynamical model is described to represent moving spots in video-microscopy image sequences. In Section V, several experimental results are reported to demonstrate the potential of the proposed approach.

II. SIMULATION FRAMEWORK AND PROPERTIES

In this section, we discuss the properties that a simulation method should respect:

- 1) Two computational approaches can be proposed for simulation: *data-driven modeling* and *physically-based modeling*. The physics-based approach exploits the physical properties of the scene and the optical characteristics of the imaging system for image modeling. The main advantage is that the model parameters are given by physics. Moreover, they are easy to interpret because they directly correspond to the real world. In return, the complexity of scenes and models is usually an obstacle to this approach and the *inverse problem* cannot be easily solved. The *data-driven modeling* aims at

describing the image sequence through statistical models learned from real images [7]. This approach can mimic dynamical processes but is not always able to describe physics of real scenes. *Data-driven* and *physically-based* approaches can be also combined to model the main components of the image. In video-microscopy, these components are mainly the moving objects, background and noise.

- 2) A simulation method must be also *controllable* [8]. This means that the representation must be parsimonious, which can help in the interpretation for an expert. In most cases, the parameters are related to physical properties of the system but also to image properties such as scale or velocity of objects. By using a such representation, the simulation method becomes more interactive and allows the expert to bring some *a priori* knowledge or to plan a set of experiments by editing the simulation. For example, an expert can indicate the locations of source and destination points of the vesicles, and by varying the positions of these points, we can observe the evolution of the simulated intra-cellular traffic. Finally, the expert feedback can be used to set up a realistic simulation.

In the proposed simulation/estimation framework, we shall see that these properties are mostly satisfied.

III. DYNAMICAL BACKGROUND MODEL

In this section, we propose a statistical framework for dynamical background modeling and estimation.

A. Image model

Large structures within the cell like the Golgi apparatus appear as nearly static during the observation time interval. In the case of images showing fluorescently tagged particles, the global image intensity is proved to vary slowly along time. This can be due to several physical phenomena such as photo-bleaching or diffusion of fluorescent proteins within the cell. Therefore, it is appropriate to propose a model able to describe the slowly spatially and temporally varying background since a stationary model would be too restrictive. The modeling of more complex dynamical small objects with variable velocities will be discussed in Section IV.

First, we have conducted experiments showing that the intensity variation with respect to time can be captured by a linear model for each pixel of the image sequence, mainly because we are dealing with sequences of limited length. This crude modeling provides a compact representation of the background intensity dynamics and the background can be described by two maps corresponding to the two spatially varying parameters of the linear model obtained for each pixel. Nevertheless, the involved parameters are spatially correlated, which will be taken into account in the estimation process. Note that the proposed method could be adapted to non-linear intensity models if required.

Formally, we propose the following image sequence model for the background

$$f(x, y, t) = a(x, y) + b(x, y)t + u(x, y, t) + \epsilon(x, y, t) \quad (1)$$

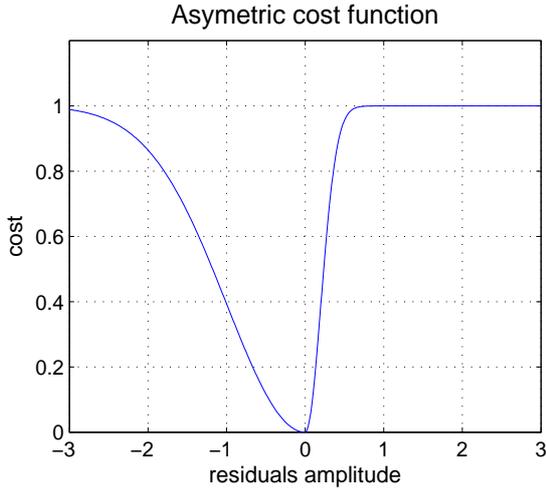


Fig. 1. Asymmetric Leclerc robust function.

where $f(x, y, t)$ denotes the intensity observed at pixel (x, y) and time t , the two coefficients $a(x, y)$ and $b(x, y)$ varies with the spatial image position $p_i = (x, y)^T$. The function $u(x, y, t)$ is a positive function that describes the intensity of moving vesicles if any, and $\epsilon(x, y, t)$ is an additive white Gaussian noise. In the sequel, we will use the subscript i to denote the spatial position p_i , and accordingly (1) can be more compactly re-written as

$$f_i(t) = a_i + b_i t + u_i(t) + \epsilon_i(t). \quad (2)$$

This model is able to describe the background intensity of the whole image sequence with only two maps $\{a_i\}$ and $\{b_i\}$ with the same size as an image of the temporal sequence. In the next section, we propose a method to estimate the maps $\{a_i\}$ and $\{b_i\}$ that describe the dynamical background model.

B. Pixel-wise estimation of the background model parameters

We first deal with the estimation of parameters a_i and b_i for a single temporal 1D signal. Let us point out that this estimation must be performed several millions of times (for each image point). Accordingly, the proposed estimation procedure must be very fast. Besides, in our target application, vesicles have an erratic behavior and sometimes stop for a long time. Consequently, prior motion detection cannot be used here to extract the objects from the background. The estimation of the dynamical background will be then based on image intensities. Also, since the background estimation must not be corrupted by the presence of moving vesicles, we will also resort to a robust estimation framework.

1) *Robust M-estimation*: The two parameters a_i and b_i are estimated by minimizing a robust error function

$$E(a_i, b_i) = \sum_{t=1}^n \rho(f_i(t) - (a_i + b_i t)), \quad (3)$$

where n is the number of samples in the 1D signal and $\rho(\cdot)$ is a robust function. A local minimum of $E(a_i, b_i)$ is commonly obtained by using the iteratively re-weighted least squares (IRLS) procedure.

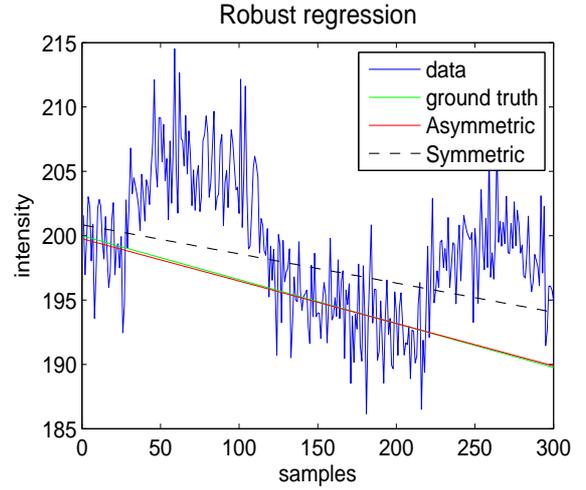


Fig. 2. Regression using an asymmetric and symmetric robust Leclerc function. The asymmetric estimator (red) fits perfectly the ground truth (green) while the symmetric function provides biased results (dotted line).

The choice of the robust function ρ is usually guided by the noise probability density function [9]. In our case, the overall noise is the sum of two components $u_i(t)$ and $\epsilon_i(t)$. In order to take into account that $u_i(t)$ usually takes high positive values (vesicles appear as bright spots in the image), we choose an asymmetric robust function (Leclerc estimator [10], [11]) plotted in Figure 1 and defined as

$$\rho(z) = \begin{cases} 1 - \exp\left(-\frac{z^2}{\lambda^2 \sigma_1^2}\right) & \text{if } z \leq 0, \\ 1 - \exp\left(-\frac{z^2}{\lambda^2 \sigma_2^2}\right) & \text{otherwise.} \end{cases} \quad (4)$$

The scale σ_2 factor can be estimated by applying a robust least-trimmed squares (LTS) estimator to the *pseudo-residuals* defined as [12] : $s_i(t) = (f_i(t+1) - f_i(t))/\sqrt{2}$, where the coefficient $1/\sqrt{2}$ ensures that $\mathbb{E}[(s_i(t))^2] = \mathbb{E}[(f_i(t))^2]$. The scale factor σ_1 is estimated by using the variance of the residuals given by the least-mean squares estimator and obtained at the initialization. Let us point out that, in regions where there are no moving vesicles, σ_1 and σ_2 are found almost equal. Finally, λ is chosen as usually in the range [1, 3].

As a matter of fact, the proposed estimator is biased [11] but the bias is small. Simulations proved that the L_2 risk of the estimator is smaller when an asymmetric cost function is used and when the data are corrupted by an additive positive signal. Figure 2 shows that the proposed estimator is able to deal with heavily contaminated data and outperforms the symmetric Leclerc M-estimator.

2) *Confidence matrix*: An accurate estimation of the confidence matrix for the estimated parameters is needed for the subsequent steps described in Section III-C. We use the approximation proposed in [9] to compute the estimation

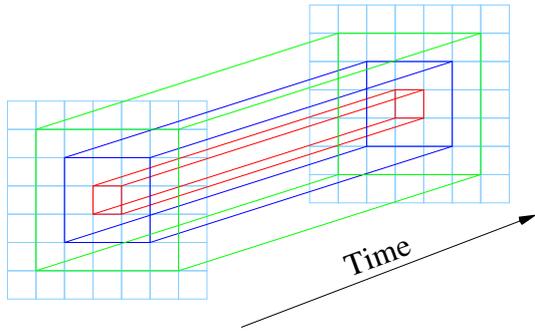


Fig. 3. Set of nested tubes $\{\mathcal{T}_{i,k}\}_{k=1,\dots,3}$.

covariance matrix defined as

$$\hat{C}_i = \frac{\sum_{t=1}^n w_{i,t}(r_i(t))r_i(t)^2 \sum_{t=1}^n w_{i,t}^2(r_i(t))}{\left(\sum_{t=1}^n w_{i,t}(r_i(t))\right)^2} \times \begin{pmatrix} \sum_{t=1}^n w_{i,t}(r_i(t)) & \sum_{t=1}^n w_{i,t}(r_i(t))t \\ \sum_{t=1}^n w_{i,t}(r_i(t))t & \sum_{t=1}^n w_{i,t}(r_i(t))t^2 \end{pmatrix}^{-1} \quad (5)$$

where $r_i(t) = f_i(t) - (a_i + b_i t)$ and the weights are given by $w(z) = \rho'(z)/z$. Unlike the expression given in [13], the approximation given by (5) is not asymptotic and yields a better estimation of the covariance matrix when n is small.

C. Spatial coherence for background estimation

We now introduce spatial coherence to regularize the maps $\{a_i\}$ and $\{b_i\}$. This can be accomplished by adopting the *bias-variance trade-off* framework described in [14], [15], [16], [17]. Instead of using a single temporal signal for each pixel p_i to estimate a_i and b_i , a set of temporal 1D signals is first collected in a spatial neighborhood of the pixel p_i . This collection of signals is then analyzed in order to take into account the desired spatial coherence of the parameters. In practice, a set of nested space-time tubes is considered (see Figure 3) by taking the pixels in a growing spatial square neighborhood of p_i . Each tube $\mathcal{T}_{i,k}$ at p_i can be parametrized by its diameter $\phi_{i,k}$ where $k \in [1, \dots, K]$ denotes the iteration:

$$\mathcal{T}_{i,k} = \{f(x_j, y_j, z_j) : |x_i - x_j| + |y_i - y_j| < \phi_{i,k}\}. \quad (6)$$

In order to select the optimal diameter of the space-time tube, we propose to minimize the point-wise L_2 risk of the parametric estimator defined as $\mathbb{E}[(\hat{\theta}_i - \theta_i)^2]$ where $\theta_i = (a_i, b_i)$ is the true parameter pair and $\hat{\theta}_i$ its corresponding estimator, at position p_i . The L_2 risk can be decomposed into two parts: squared bias and variance. As shown in Figure 4, while the diameter $\phi_{i,k}$ increases with k , the bias increases too. This can be explained by the fact that, the data cannot

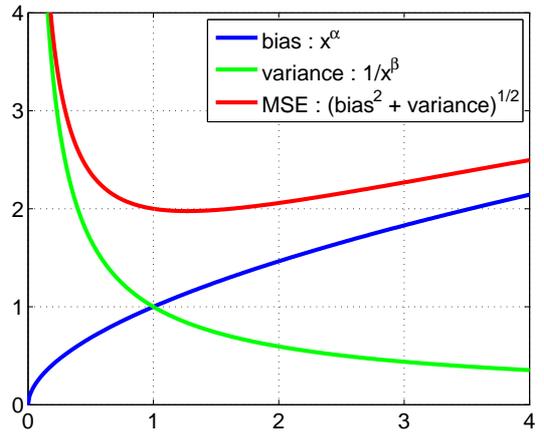


Fig. 4. *Bias-variance trade-off* principle. When the diameter of the tube increases, the bias increases and the variance decreases. The optimum is achieved when the bias and the variance are of the same order.

be described any longer by a unique model. In contrast, by taking more and more data points, the variance decreases. This behavior, also called *bias-variance trade-off* [14], is exploited to detect the minimum of the point-wise L_2 risk which is nearly equal to twice the variance (see Fig. 4).

For each diameter $\phi_{i,k}$, new estimates of the background model parameters $\hat{\theta}_{i,k}$ and the associated covariance matrix $\hat{C}_{i,k}$ are computed with the same procedure as the one described in Section III-B but using all the data taken in the considered spatial neighborhood. It can be shown that the *bias-variance trade-off* can be expressed with the following test [17] (here a vectorial version is given):

$$\frac{n-2+1}{2n} \left(\hat{\theta}_{i,k} - \hat{\theta}_{i,k'}\right)^T \hat{C}_{i,k'}^{-1} \left(\hat{\theta}_{i,k} - \hat{\theta}_{i,k'}\right) < \eta \quad (7)$$

for all $1 \leq k' < k$. While this inequality is satisfied, the diameter of the tube is increased and the estimation process is continued. It can be proved that the threshold η can be defined as a quantile of a Fisher distribution of parameters 2 and $n-2-1$.

In this section, we have proposed a spatially and temporally varying background and a statistical estimation framework to estimate the involved parameters in the modeling. In the second part of the paper, we propose a simulation framework to produce dynamics corresponding to moving small spots in the image sequences.

IV. SPOT MODEL

In video-microscopy, vesicles appear in many image sequences as small bright spots. The object diameters theoretically ranges from 60 nm to 150 nm. The resolution of the microscope is about $130 \times 130 \times 300$ nm. Then, the diameters of spots are often below this resolution. However, the point spread function of the video-microscope make them appear as larger structures even if a deconvolution process [18] is applied. Furthermore, when the density of objects increases, vesicles gather together and constitute small rods.

These vesicles are also known to move along microtubules, that is along long protein polymers that have an exceptional bending stiffness and can be easily fit by smooth curves. Microtubules are conveyor belts inside the cells. They drive vesicles, granules, organelles like mitochondria, and chromosomes via special attachment proteins using molecular motors. It is also established that molecular motors form a class of proteins responsible for the intra-cellular transport within the cells. The dynein and kinesin proteins are two classes of motors associated with microtubules. It has been shown that the concentration of these molecular motors influences the structure and the dynamics of the microtubule network. In stable conditions, the speed of these motors is assumed to be constant. This explains why the observed velocity of vesicles is constant if they move along the same microtubule.

In our target application, vesicles move along the microtubule, leaving a donor organelle and reaching an acceptor organelle, e.g. the Golgi apparatus to the Endoplasmic Reticulum. In the proposed simulation method, we do not simulate the dynamics of the microtubules like in [2] but we rather aim at describing video-microscopy image sequences using a data-driven approach.

A. Photometric model

In most cases, vesicles are represented by anisotropic Gaussian spots with variances related to the spot size ranging from 60 nm to 150 nm. Then, the size of the vesicles in the image is close to the pixel size. The size of the spots will be estimated on image sequence as described in the Section IV-C.

Now, the covariance matrix is a function of the velocity direction in order to get the major axis of the ellipse corresponding to the displacement of the vesicle. The ellipticity is also a function of the velocity. In Fig. 5, we can see how the covariance matrix of the Gaussian function allows to modify the orientation of the spot according to the direction of the microtubule.

In addition, the spots can merge and split and the ellipticity of the spots allows to easily simulate rods. These rods are made of a collection of spots which use the same path. In the resulting image sequence they seem to describe a single object since they have the same velocity while they are on the same microtubule.

B. Dynamical model

1) *Network modeling*: A physics-based simulation of the self-organization of the microtubule network can be found in [19]. It is based on the interaction between the motors (e.g. kinesin) and microtubules, and explains some characteristic conformations such as mitotic spindle. It takes into account the dynamical behavior of the microtubules. However, this computer simulation only describes the behavior of the microtubule network *in-vitro* and is not adapted for the more complex *in-vivo* case in which the microtubules interact with other organelles of the cell. In addition, the observation time intervals are usually short compared to the dynamics of the network itself. Accordingly, we propose to adopt a static model for the network.

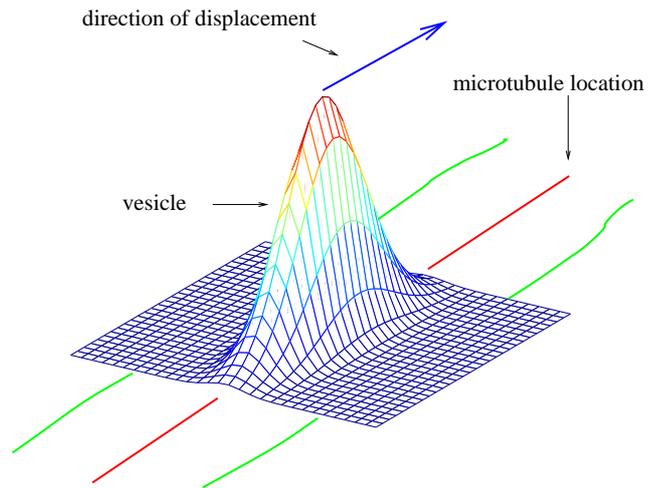


Fig. 5. Gaussian spot oriented in the direction of the microtubule. The covariance matrix of the Gaussian function is defined as a function of the velocity of the vesicle and the simulated vesicles are then elongated along with the motion direction.

In order to produce a synthetic but realistic microtubule network, we exploit real image sequences as inputs for the modeling. The network could be tagged with Green Fluorescence Protein (GFP) but this network is too complex and individual microtubules cannot be easily extracted. However, the microtubule network can be also crudely computed from a maximum intensity projection map *wrt* time, that is from the paths used by the tagged objects. Figure 6 shows the maximum intensity projection map of a sequence made of 300 images. This simple method allows to select a subset of paths mainly used for the intra-cellular trafficking leading to network with low complexity; this approach has been successfully used for the construction of kymograms [20]. However, as shown in Fig. 6, all the paths are not complete, especially if the sequence duration is too short. The gaps are then completed by using a painting software. Furthermore, the positions of the roads are extracted from the network image using the *unbiased line detection* algorithm described in [21]. Finally, each road is finally described by its length, its width and its source node and its destination node.

2) *Selection of source/destination nodes*: In the proposed simulation, vesicles are going from one point to another. Typically, they leave a donor organelle and move toward an acceptor organelle. Once the network has been computed, the expert needs to specify the source and destination nodes on the network. In order to take into account the lack of *priori* information on the organelle and their function, a node can be both a source and a destination while the other nodes represent the intersection points of the network and are only used for the routing. Source-destination pairs are important cues for the simulation and corresponds to a birth/death map as described in [8]. These labels are actually related to the locations and relationships of specific organelles within the cell.

In Fig. 7, the source and destination nodes have been manually selected. The destination nodes labeled in red correspond

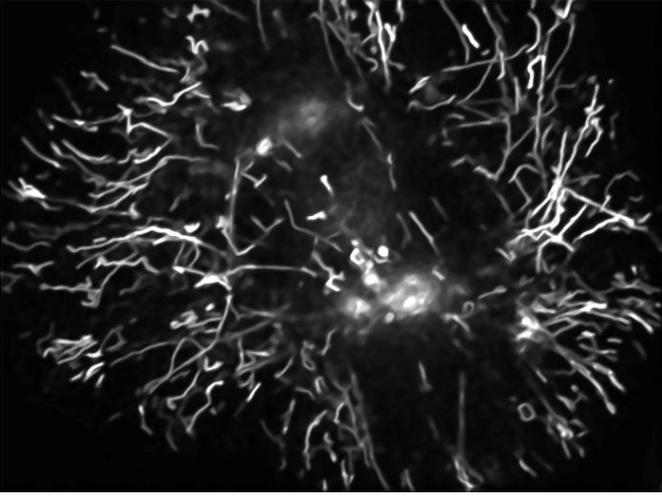


Fig. 6. Maximum intensity projection map computed from an image sequence. The paths used by the vesicles appear as bright filaments. The maximum intensity projection map has been simplified using the algorithm described in [17].

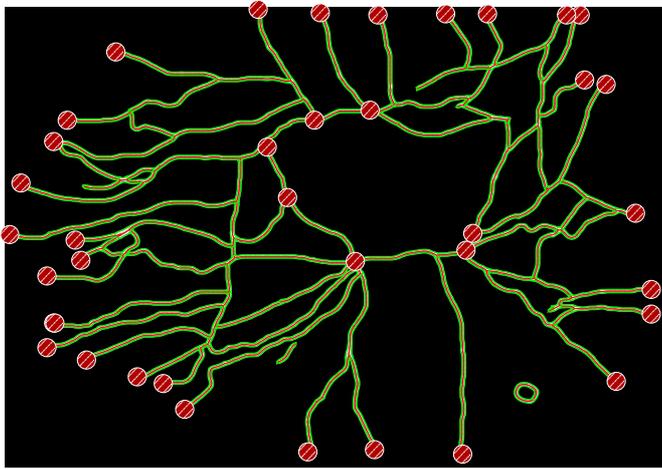


Fig. 7. Representation of a realistic synthetic network. This network is based on a maximum intensity projection map and has been manually simplified. This network is composed of 146 nodes and 160 bi-directional links which correspond to 320 directional edge in the graph associated to the network.

to end-points while the source nodes corresponding to the membrane of the Golgi apparatus are labeled in green. In this simulation, vesicles are only going from the Golgi to the end-points located at the periphery of the cell. Thus, the retrograde transport from end-point to the Golgi is prohibited for simplicity.

In our approach, the paths defined as the minimal paths between the source and the destination nodes are computed by using the Dijkstra algorithm. In that case, the weight associated to each edge can be defined as a function of the length of the corresponding road but it can also take into account other parameters. Note that the speed associated to edges can be also used to estimate the shortest path. Finally, as expected, the vesicles move along the estimated roads with velocities given by the speed-limits of the roads. At each time step, the vesicle is then displaced along the microtubule with a distance

which is proportional to the velocity.

C. Estimation of model parameters and spot detection

Object detection in image sequences is an important task in video-microscopy but generally ground truth is not always available. In this section, we propose a method to decide which pixels belong to the image background. Furthermore, the unsupervised approach is applied to really and artificially image sequences obtained from the modeling framework described in the paper.

In our detection method, a penalized likelihood criterion is introduced to estimate the mean and the variance of a Gaussian vector and by assuming some of the components are zero. We suppose that the number of zero components in the 1D temporal signal as well and their time positions are unknown. In our problem, the zero components of the 1D signal for position p_i are associated to the background while the other components are related to the moving objects (see Fig. 8). Formally, we define

$$\mathbf{Z} = \mathbf{m} + \mathbf{e}, \quad \mathbf{e} \sim \mathcal{N}(0, \Sigma), \quad (8)$$

where \mathbf{Z} is the vector of n residuals given by $z(t) = f_i(t) - (a_i + b_i t)$, assumed to be independent, and \mathbf{e} is a Gaussian white noise with variance $\tau^2 I_n$. The expectation $\mathbb{E}[\mathbf{Z}] = \mathbf{m} = (m_1, \dots, m_n)^T$ and the variance τ^2 are unknown. Moreover, k_0 components of \mathbf{m} are assumed to be non zero. The data assumed to be independent, are then re-ordered such that $|z_j| > |z_{j+1}|$. Now, we consider the following model collection $\mathcal{M} = \{M_1, \dots, M_n\}$

$$\begin{cases} M_1 = (z_1, 0, \dots, 0)^T \\ M_2 = (z_1, z_2, 0, \dots, 0)^T \\ M_n = (z_1, z_2, z_3, \dots, z_n)^T. \end{cases} \quad (9)$$

The selected model corresponding to object/background labeling, is the one that minimizes the following penalized likelihood criterion [22], [23]

$$J(M_j) = \frac{n}{2} \log(\hat{\tau}_j^2) + n \left(c_1 \log \left(\frac{n}{k_j} \right) + c_2 \right) \frac{k_j}{n - k_j} \quad (10)$$

where $\hat{\tau}_j^2 = n^{-1} \sum_{k=j}^n z_k^2$ is the plugged-in maximum likelihood estimator of τ_j^2 . The two universal constants c_1 and c_2 were calibrated in [23] and found to be $c_1 = 2$ and $c_2 = 4$.

Once a model $M_{k_0}(p_i)$ has been estimated, which amounts to select an optimal threshold $t_{k_0}(p_i)$ for each point p_i in the image, a minimal threshold is t_{min} calculated in order to take into account the fact that no object passed through some points of the image. We decide that these points belong to the background during the whole image sequence (see Fig. 8). Then, we apply the same thresholding method based on penalized criterion to the spatial threshold map. Thus, this method provides a set of points belonging to the background, and by taking the complementary set, we detect the moving vesicles.

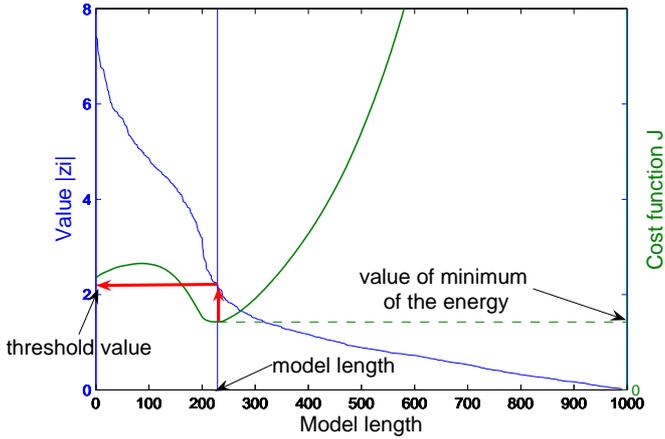


Fig. 8. Illustration of the *selection model* principle applied to a synthetic signal ($n = 1000$). The number of samples corresponding to the background is about 800. The X-axis represents the length of the model M_j ; the left ordinate corresponds to the value $|z_j|$ associated to the model M_j labeled in blue; the right ordinate represents the value of the cost function $J(M_j)$.

V. EXPERIMENTS

In this section, we describe two experiments that demonstrate the potential of the proposed simulation method. First, a simulation of a synthetic image sequence based on real image sequence is described. Besides, we describe a synthetic image sequence whose parameters have been manually determined to test a denoising procedure as well as the estimation procedure for the background model described in Section III.

A. Realistic image sequence simulation

We first propose to use a real image sequence in order to simulate a sequence with similar photometric and dynamical contents. The original sequence is shown in Fig. 9(a) and represents vesicles moving from the Golgi apparatus to the Endoplasmic Reticulum. The parameters of the dynamical background are first estimated as described in Section III. The two maps $\{a_i\}$ and $\{b_i\}$ are respectively shown in Fig. 9(e) and in Fig. 9(f). Once these parameter maps have been estimated, the background is subtracted from the original image in order to obtain the sequence of residuals shown in Fig. 9(g), which is a noisy representation of the moving spots. The paths mainly used by vesicles during the 150 frames of the real sequence can be observed on the maximum intensity projection map in the time direction, of residuals as shown in Fig. 9(b). We propose also to enhance the maximum intensity projection map using optimal steerable filters [24] (see Fig. 9(c)). The *unbiased line detection* algorithm [21] is then applied to the enhanced image in order to estimate the positions of the roads shown in Fig. 9(d). Finally 150 vesicles are generated and moved along the estimated network. The velocities of the vesicles are tuned so that the simulated sequence provides the same visual effect than the original sequence. It is confirmed by experts that the proposed simulation method provides a very realistic image sequence both from the photometric point of view as well as from the dynamical point of view.

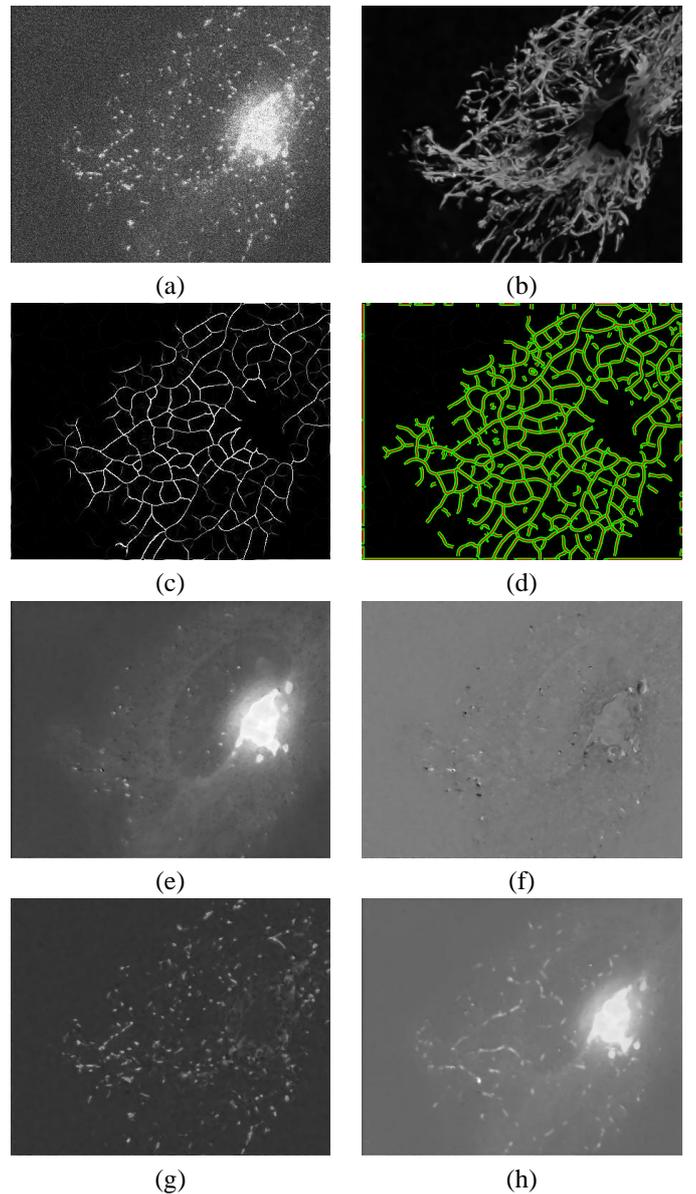


Fig. 9. Simulation of a video-microscopy image sequence from a real image sequence. (a) one frame of the maximum intensity projection sequence *wrt* z axis (depth) computed from an original $3D+time$ image sequence; (b) maximum intensity projection $2D$ map *wrt* time t ; (c) results of steerable filtering; (d) results of the un-biased line detection; (e) map $\{a_i\}$; (f) map $\{b_i\}$; (g) residual map; (h) image reconstruction from estimated parameters.

B. Benchmarks

In order to evaluate the quality of the proposed background estimation method, we have simulated a $128 \times 128 \times 150$ image sequence. This simulation is composed of a 20 vesicles moving along a network shown in Fig. 10(a). The two maps $\{a_i\}$ and $\{b_i\}$ for the background model are shown in Figs. 10(b)-(c). By adding to these three images, the photometric characteristics of the spots, we get the set of parameters that controls the simulation.

To generate the background model, we have manually designed the shape of the background shown in Fig. 10(b). Then, we have computed the image shown in Fig. 10(c) for a duration equal to twice the duration of the simulation, to

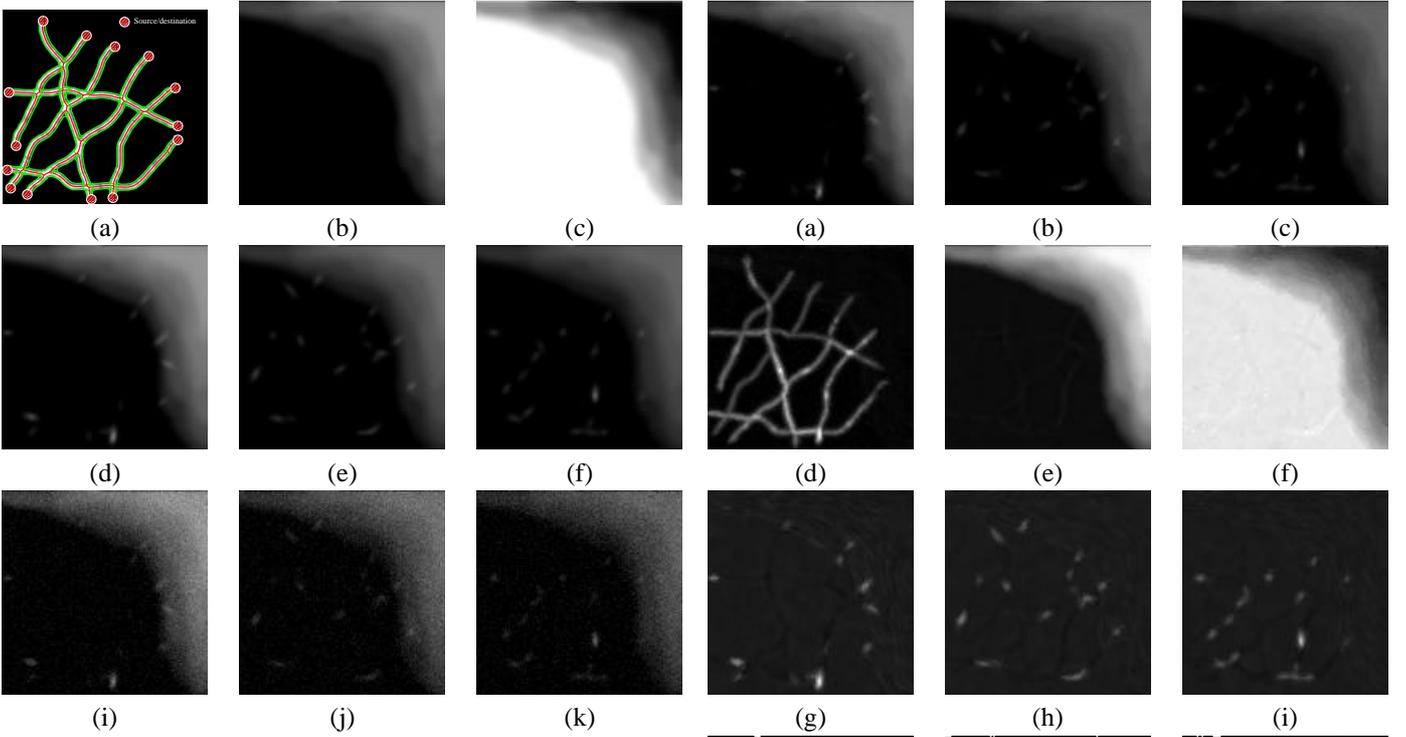


Fig. 10. Simulation of a synthetic video-microscopy image sequence. (a) network map; (b) map $\{a_i\}$ corresponding to the background; (c) map $\{b_i\}$ corresponding to the variation of the background; (d,e,f) three frames extracted from the noise-free synthetic image sequence; (i,j,k) three noisy frames corresponding to a signal-to-noise ratio of $13.6dB$.

obtain a uniform background. Three frames of the noise free-simulated sequence are shown in Fig. 10(d)-(f). A Gaussian noise of standard deviation $\sigma = 9$ has been also added to these frames and the results are shown in Fig. 10(i,j,k). The intensity of the vesicles are assumed to follow a Gaussian distribution with mean 30 and standard deviation 3.

Now, we first apply the algorithm described in [25] to denoise the image sequence artificially corrupted by a Gaussian white noise of variance 9 which correspond to a signal-to-noise of $13.6dB$. By using the pointwise adaptive space-based $3D\ 3 \times 3 \times$ patch based approach with 6 iterations, the mean squared error of the recovered image is finally 1.35, which corresponds to a signal-to-noise ratio of $30.0dB$ (decibels). The noise is then drastically reduced (see Fig. 11) and, visually the sequence looks similar to the original artificial sequence and no vesicle has disappeared. Note that denoising algorithms are commonly applied to image sequences in video-microscopy as a pre-processing step in practical imaging.

In order to estimate the performance of the proposed estimation method for the dynamical background model, we have applied the estimation procedure in order to recover the simulation parameters. The two estimated maps $\{a_i\}$ and $\{b_i\}$ are shown in Fig. 11. The mean squared error between the original map $\{a_i^o\}$ and the estimated map $\{\hat{a}_i\}$ is 1.50 and the signal to noise ratio is $31.44dB$. The signal-to-noise ratio calculated from the original map $\{b_i^o\}$ and the estimated map $\{\hat{b}_i\}$ is $32.70dB$.

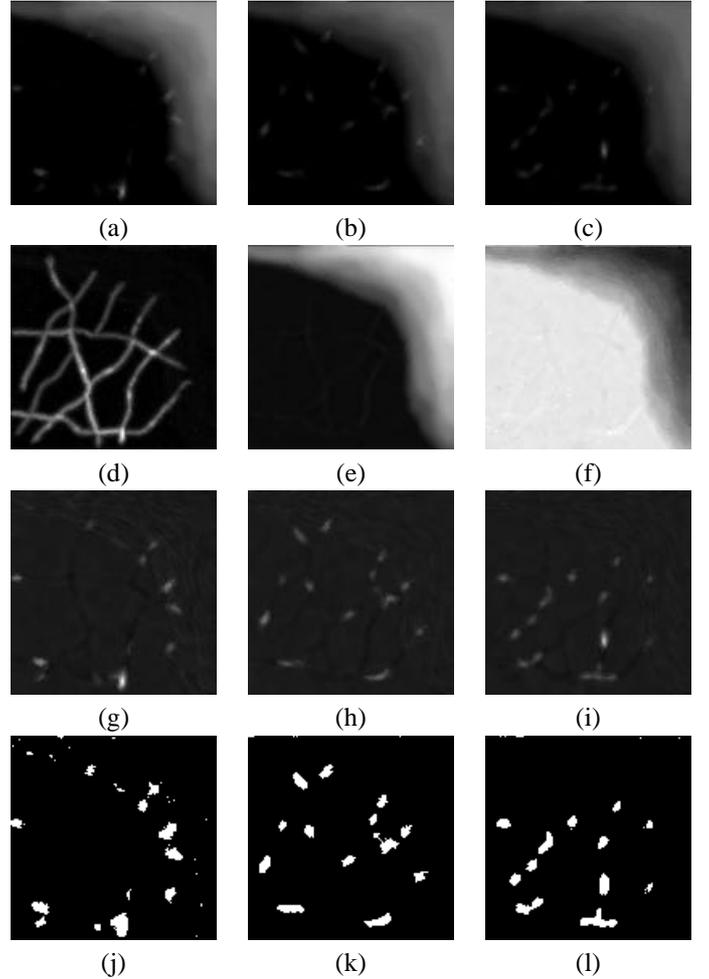


Fig. 11. Estimation of a synthetic video-microscopy image sequence. (a,b,c) three frames of the denoised image sequence; (d) maximum intensity projection map; (e) estimated map $\{\hat{a}_i\}$; (f) the estimated map $\{\hat{b}_i\}$; (g,h,i) residual maps after background subtraction; (j,k,l) detection maps using the described method.

VI. CONCLUSION

In video-microscopy, tracking methods that estimate trajectories of small objects (particles) may encounter difficulties if the number of objects is large and the signal-to-noise ratio is low. Moreover, the tracked objects are not always visible in the sequence and data association is problematic. To evaluate the performance of the object tracking algorithms in video-microscopy, we have presented a simulation/estimation framework to artificially produce image sequences that can mimic mostly moving spots observed in video-microscopy. This new computational tool can be also used by an expert for inspecting real image data. In that case, the user adapts the simulation parameters to the observed data, and then selects a set of parameters that can be considered as a parsimonious representation of the image sequence contents. For future works, we plan to better validate this approach in collaboration with expert-biologists.

REFERENCES

- [1] F. Gibbons, J. Chauwin, M. Despósito, and J. José, "A dynamical model of kinesin-microtubule motility assays," *Biophysical Journal*, vol. 80,

- no. 6, pp. 2515–2526, June 2001.
- [2] F. Nédélec, “Computer simulations reveal motor properties generating stable antiparallel microtubule interactions,” *Journal of Cell Biology*, vol. 158, no. 6, pp. 1005–1015, Sept. 2001.
 - [3] A. Genovesio, T. Liedl, V. Emiliani, W. Parak, M. Coppey-Moisan, and J.-C. Olivo-Marin, “Multiple particle tracking in 3D+t microscopy: Method and application to the tracking of endocytosed quantum dots,” *IEEE Trans. on Image Processing*, vol. 15, no. 5, pp. 1062–1070, 2006.
 - [4] V. Racine, A. Hertzog, J. Jouanneau, J. Salamero, C. Kervrann, and J. Sibarita, “Multiple target tracking of 3d fluorescent objects based on simulated annealing,” in *Proc. of IEEE International Symposium on Biomedical Imaging, ISBI’2006*, Apr. 2006.
 - [5] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, “Image denoising using scale mixtures of gaussians in the wavelet domain,” *IEEE Trans. on Image Processing*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.
 - [6] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, “Performance of optical flow techniques,” *International Journal of Computer Vision*, vol. 12, no. 1, pp. 43–77, 1994.
 - [7] S. Soatto, G. Doretto, and Y. N. Wu, “Dynamic textures,” in *Proc. of 8th IEEE Int. Conf. on Computer Vision, ICCV’2001*, vol. 2, Vancouver, July 2001, pp. 439–446.
 - [8] Y. Wang and S. Zhu, “Modeling textured motion: Particle, wave and sketch,” in *Proc. of 9th IEEE Int. Conf. on Computer Vision, ICCV’2003*, 2003.
 - [9] S.-S. Ieng, J.-P. Tarel, and P. Charbonnier, “Evaluation of robust fitting based detection,” in *Proc. of 8th Eur. Conf. on Computer Vision, ECCV’2004*, vol. 2, Prague, May 2004, pp. 341–352.
 - [10] H. Allende, A. Frery, J. Galbiati, and L. Pizarro, “M-estimator with asymmetric influence function: the \mathcal{G}_a^0 distribution case,” *Journal of Statistical Computation and Simulation*, in press.
 - [11] A. Ruckstuhl, M. Jacobson, R. Field, and J. Dodd, “Baseline subtraction using robust local regression estimation,” *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 68, no. 2, pp. 179–193, Jan. 2001.
 - [12] T. Gasser, L. Sroka, and C. Jennen Steinmetz, “Residual variance and residual pattern in nonlinear regression,” *Biometrika*, vol. 73, pp. 625–633, 1986.
 - [13] P. Huber, *Robust Statistics*. John Wiley and Sons, New York, NY, 1981.
 - [14] O. Lepski, “Asymptotically minimax adaptive estimation 1: upper bounds,” *SIAM Journal Theory of Probability and Application*, vol. 36, no. 4, pp. 654–659, 1991.
 - [15] M. Maurizot, P. Boutheymy, B. Delyon, A. Juditski, and J. Odobez, “Determination of singular points in 2D deformable flow fields,” in *Proc. of IEEE Int. Conf. on Image Processing, ICIP’1995*, vol. 3, Washington DC, Oct. 1995, pp. 488–491.
 - [16] C. Ercole, A. Foi, V. Katkovnik, and K. Egiazarian, “Spatio-temporal pointwise adaptive denoising in video: 3d non parametric approach,” in *Proc. of the 1st International Workshop on Video Processing and Quality Metrics for Consumer Electronics, VPQM’2005*, Scottsdale, 2005.
 - [17] C. Kervrann and J. Boulanger, “Optimal spatial adaptation for patch-based image denoising,” *IEEE Trans. on Image Processing*, to appear.
 - [18] J.-B. Sibarita, H. Magnin, and J. De Mey, “Ultra-fast 4D microscopy and high throughput distributed deconvolution,” in *Proc. of IEEE Int. Symp. on Biomedical Imaging, ISBI’2002: Macro to Nano*, Washington, United States, 2002, pp. 769–772.
 - [19] T. Surrey, F. Nédélec, S. Leibler, and E. Karenti, “Physical properties determining self-organization of motors and microtubules,” *Science*, vol. 292, pp. 1167–1171, May 2001.
 - [20] J. Sibarita, V. Racine, and J. Salamero, “Quantification of membrane trafficking on a 3d cytoskeleton network in living cells,” in *Proc. of IEEE International Symposium on Biomedical Imaging, ISBI’2006*, Apr. 2006.
 - [21] C. Steger, “An unbiased detector of curvilinear structures,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 2, pp. 113–125, 1998.
 - [22] L. Birgé and P. Massart, “Gaussian model selection,” *Journal of European Mathematical Society*, vol. 3, pp. 203–268, 2001.
 - [23] S. Huet, “Model selection for estimating the non zero component of a gaussian vector,” University of Paris-Sud, Math. Dept., Tech. Rep. 25, 2002.
 - [24] M. Jacob and M. Unser, “Design of steerable filters for feature detection using canny-like criterion,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 8, pp. 1007–1019, Aug. 2004.
 - [25] J. Boulanger, C. Kervrann, and P. Boutheymy, “Space-time adaptation for patch-based image sequence restoration,” IRISA, Rennes, France, Tech. Rep. PI 1798, Apr. 2006.