Advances in probabilistic particle tracking for biomedical imaging

Nicolas Chenouard

To cite this version:

Nicolas Chenouard. Advances in probabilistic particle tracking for biomedical imaging. Signal and Image processing. Télécom ParisTech, 2010. English. NNT : . tel-00560530

Institut Pasteur - TELECOM ParisTech

Ecole Doctorale d'Informatique, Télécommunications et
Electronique de Paris

Dissertation

presented in fulfillment of the requirements for the degree of
Doctor of Philosophy of TELECOM ParisTech

Domain : Signal and Image Processing

by Nicolas Chenouard

Advances in probabilistic particle tracking for biological imaging

Defensed on January 21, 2010

Jury

Jean-Christophe	OLIVO-MARIN	Advisor
Isabelle	BLOCH	Co-Advisor
Erik	MEIJERING	Reviewer
Patrick	PÉREZ	Reviewer
Laurent	COHEN	President
Benoît	DUBERTRET	Examiner

Institut Pasteur Paris
Unité d'Analyse d'Images Quantitative
Département de Biologie Cellulaire et
Infection

TELECOM ParisTech
Laboratoire Traitement et
Communication de l'Information
Département Traitement du Signal et
des Images

Acknowledgments

First, I would like to express my sincere gratitude to Jean-Christophe Olivo-Marin, my advisor and head of the Quantitative Image Analysis Unit of Institut Pasteur. His guidance and his trust in my work along my Ph.D. project have made the present thesis possible.
I am deeply grateful to my co-advisor, Isabelle Bloch, professor at the department of Signal and Image Processing at Télécom ParisTech. She has given me her valuable comments, suggestions, criticisms, and has offered me a continuous and precious encouragement.

My sincere thanks are due to the jury members, Laurent Cohen, Professor at CEREMADE, University Paris IX Dauphine, Erik Meijering, Professor at the Medical Center Erasmus in Rotterdam, Patrick Pérez, directeur de recherche at Thompson Technology Rennes Laboratory, Benoît Dubertret, chargé de recherche at ESPCI ParisTech, who have reviewed my work, given me their advices, criticisms and constructive comments. I also thank Henri Maître, directeur de l'Ecole Doctorale Informatique, Télécommunications et Electronique de Paris, for the support of the doctoral school.

I am greatly indebted to all the collaborators who have taken part in my Ph.D. project.
I am grateful to Benoît Dubertret, Pedro Felipe Gardeazabal Rodriguez and Sergey Sosnovski from ESPCI, and to Christophe Machu from the PFID at Pasteur, for their help and the opportunity they gave me to experiment the principles of image acquisition.
I wish to express my gratitude to the biologists with whom I collaborated along my thesis : Johanna Buisson, Philippe Bastin, Karine Gousset, Chiara Zurzolo, Musa Mhlanga, Vincent Galy, all from Institut Pasteur, Emmanuelle Nicolas and Antoine Guichet from Institut Jacques Monod, Elizabeth Crowell and Samantha Vernhettes from INRA, Roli Richa and Alexandra Fragola from ESPCI. Their original biological projects have been a great source of inspiration for the development of novel mathematical tools, and they have given a real sense to my methodological work.

I am grateful to the people from the group Traitement et Interprétation des Images of Télécom ParisTech for interesting discussions, and I would like to thank all my past and present colleagues in the Quantitative Image Analysis Unit : Fabrice de Chaumont, Vannary Meas-Yedid, Marcio de Moraes Marim, Alexandre Dufour, Bo Zhang, Sylvain Berlemont, Christophe Zimmer, Clovis Tauber, Stéphane Dallongeville, François Orieux, Inès Széliga, Nicolas Hervé, for their help and warm support during my work. More specifically, I wish to thank Fabrice de Chaumont wholeheartedly for his great help during my Ph.D. project and his friendship. I am also grateful to Bo Zhang for his guidance which was inspiring to me in the first year of my project, and to Alexandre Dufour and Marcio de Moraes Marim for fruitful scientific (and unscientific !) discussions.

This research was funded by the C'Nano Ile-de-France, CNRS, and Institut Pasteur, and I acknwoledge their support.

Finally, I am deeply grateful to Lisa and my family for their continuous support, encouragement and kindness.

Abstract

Particle tracking is a method of choice to understand subcellular mechanisms since it provides a robust and accurate means to characterize the dynamics of moving objects at the micro- and nano-metric scale. This thesis addresses several aspects related to the problem of tracking several hundreds of particles in cluttered conditions. We present novel techniques based on robust mathematical methods which allow us to track subresolutive particles in the wealth of conditions which are met in cellular imaging.

Particle detection : we have first addressed the issue of detecting particles in fluorescence images containing a structured background. The key idea of the proposed method is to use a blind source separation technique : the Morphological Component Analysis (MCA) algorithm, to separate the background from the particle signal by exploiting the differences between their morphologies in the image. We have made a number of adaptations to the MCA to comply with the characteristics of biological fluorescence images. For instance, we have proposed to use the curvelet dictionary and a wavelet dictionary with sparsity priors to split the background and the particles signals. After source separation, the background-free image can be reliably analyzed to identify the particle locations and track them through time.

Particle tracking problem modeling : we have proposed a global statistical framework which accounts for every aspect of the particle tracking problem in cluttered conditions. The designed probabilistic framework includes several models which are dedicated to biological imaging, such as statistical models of motion of particles. We have also defined the concept of target perceivability for biological particles. By doing so, the existence of particles is explicitly modeled and statistically quantified, thereby addressing the issues of track termination/creation within the statistical framework for tracking. As a result, the proposed framework enjoys a high degree of flexibility since every parameter and model involved finds a simple and intuitive interpretation. The proposed probabilistic framework has consequently allowed us to exhaustively model a wealth of different biological cases.

Tracking algorithm design : we have reformulated the Multiple Hypothesis Tracking (MHT) algorithm to include the probabilistic framework for tracking biological particles, and proposed an efficient implementation which allows one to track numerous particles in poor imaging conditions. The *Enhanced* MHT (E-MHT) we propose takes full advantage of the tracking model by incorporating the knowledge from future frames, whereby the significance of statistical scores is increased. As a result, the E-MHT is able to automatically identify false detections and detect target appearance/disappearance events. We address the complexity of the tracking task by an efficient design of the algorithm which exploits the tree organization of the solutions and the ability to make the computations in a parallel manner. A series of comparative tests between the E-MHT and a number of state-of-the-art techniques have been performed with synthetic 2D image sequences and real 2D and 3D data sets. In every case the E-MHT has proved superior performance over standard techniques, with a remarkable capability to handle very poor imaging conditions.

We have applied the proposed tracking techniques in a number of biological projects, leading to novel biological results. The flexibility and robustness of the proposed methods have allowed us to track prions infecting cell, characterize protein transport during the Drosophila oocyte development, and to study the mRNA trafficking in a Drosophila ovocyte.

Contents

List of Figures

List of Tables

Introduction

1 Characterizing the dynamics of biological processes

The cellular machinery is governed by a complex interplay of highly dynamic processes at the cellular and molecular levels. Investigating these processes is thus necessary step toward the understanding of higher order biological systems such as embryogenesis, infectious diseases mechanism, neuronal communication, or the immune response to pathogenic invasion. However, the study of sub-cellular mechanisms has long been complicated by technical issues such as the nanometric scale of the biological systems and their fast dynamics which have put them out of reach of biologists eyes for years. The current decade is however seeing a major breakthrough in cellular biology with the development of new image-based analysis methods that are becoming reference tools for assessing novel biological results in a reliable and reproducible manner.

1.1 Imaging cellular processes

The path towards the understanding of dynamical processes has been significantly enlightened over the past decades thanks to the numerous advances in fluorescent probes, labeling techniques and microscopy systems : one is nowadays able to specifically label some nanometric scale objects with bright and stable fluorophores [114, 56] which can be observed with fast and sensitive cameras thanks to microscopes achieving nanometric accuracy [83, 142, 17]. Imaging techniques thus allow biologists to look directly at the biological system at the appropriate scale both in space and time. It is now possible to look in real-time at viruses or prions infecting a cell [6, 78], as well the trafficking of proteins between cells [126]. Fluorescence imaging protocols have consequently led scientists to shed light on a large amount of biological phenomena both at the cellular and sub-cellular levels, many of which had never been observed in other imaging modalities [179]. Some dedicated analysis techniques to these new types of data are however required. The need for well suited methods is emphasized by the intricate link between the study of a dynamic process and its spatio-temporal scale, both on the imaging and on the quantification sides. More specifically, the distinction between cellular and molecular scales has led to different imaging techniques and different approaches to analyze resulting data. Systematic image acquisition and quantification of biological processes yield numerous issues at each of these scales, hence we refer to [189] and references therein for studies on cell dynamics, while our work focuses on the sub-cellular mechanisms, such as virus invasion and protein exocytosis, which involve highly dynamic nanometric scale objects and fast interactions.

1.2 Motion analysis of single particles at nano-metric scale

The characterization of fast nanometric scale processes is of major interest to understand sub-cellular and cellular mechanisms. For instance, viral infection processes include many steps involving dynamic interactions with different cellular structures such as the binding of the virus to specific receptors or attachment factors on the cell surface, and the transport to specific sites mediated by the cell cytoskeleton. Small particles may exhibit rapid motion within cells, especially when supported by an active transport, e.g., mediated through F-actin, myosin or microtubules networks [109]. For instance, viruses can spread from cell to cell on small protrusions either driven by actin polymerization or along thin connecting fibers [28]. The diversity of these interactions, and the possible existence of alternative infection routes, make it necessary to study individual viral particles paths. As a solution, single particle tracking allows one to visualize and characterize the path of dynamic nanometric objects, and provides a straightforward and accurate means to understand pathogenic mechanisms. Single-virus tracking has consequently emerged over the past few years as a powerful technique for the study of virus trafficking (see [28] and references therein for a comprehensive review) or prions spreading on tunneling nanotubes [78].

Particle tracking has become a popular tool in a number of other fields of biology, such as the study of membranes structuration and dynamics [152, 146]. Over the past decades, most experimental measurements of the lateral diffusion of membrane lipids and proteins were carried out by fluorescence photobleaching recovery (FPR, read [58] and references therein), which aims at inferring particles motility characteristics from the observation of the fluorescence recovery in a bleached region, as opposed to the study of individual particle movements. The FPR method has however been proved inefficient for discriminating sub-groups of particles with little motility differences [4]. In contrast with such a global analysis, single particle tracking provides a direct measurement of dynamical processes, allowing an exhaustive characterization.

The advent of particle tracking as a reference tool in biology yields new challenges on the image analysis side. The vast amount of data produced for particle tracking experiments and the need for reliable and reproducible results yield the requirement of novel methods able to automatically extract particle trajectories in fluorescence image sequences.

2 Particle tracking in biology

2.1 Common problems

When imaging cellular processes, fluorescently labeled particles appear as compact and bright objects which we call *spots*. An intuitive idea to build the trajectories of the spots over a sequence of images is to proceed in two steps :

1. detect the spot positions in each frame,
2. link the sets of positions between subsequent frames according to the supposed particles.

We show in Figure 1 the main principles of this approach for multiple particle tracking when the procedure is applied sequentially to a set of images. A wealth of tracking solutions have been developed relying on this simple scheme and they are nowadays found in most microscopy images analysis softwares. Satisfactory results may be obtained when the processed image sequence shows a limited number of objects that are easily identifiable and move slowly. However, numerous complicating factors have to be dealt with in most biological applications, such as a high density of targets and their fast and changing dynamics. Designing advanced image processing tools is thus required to solve the issues inherent to the imaging of sub-cellular dynamic processes. The standard two-step procedure for tracking (Fig. 1) suffers from important limitations in the vast majority of biological applications. First, the detection of spots rarely succeeds in identifying accurately the position of all particles. Small particles may indeed have a low intensity in fluorescence images, which make them difficult to discriminate from the background.

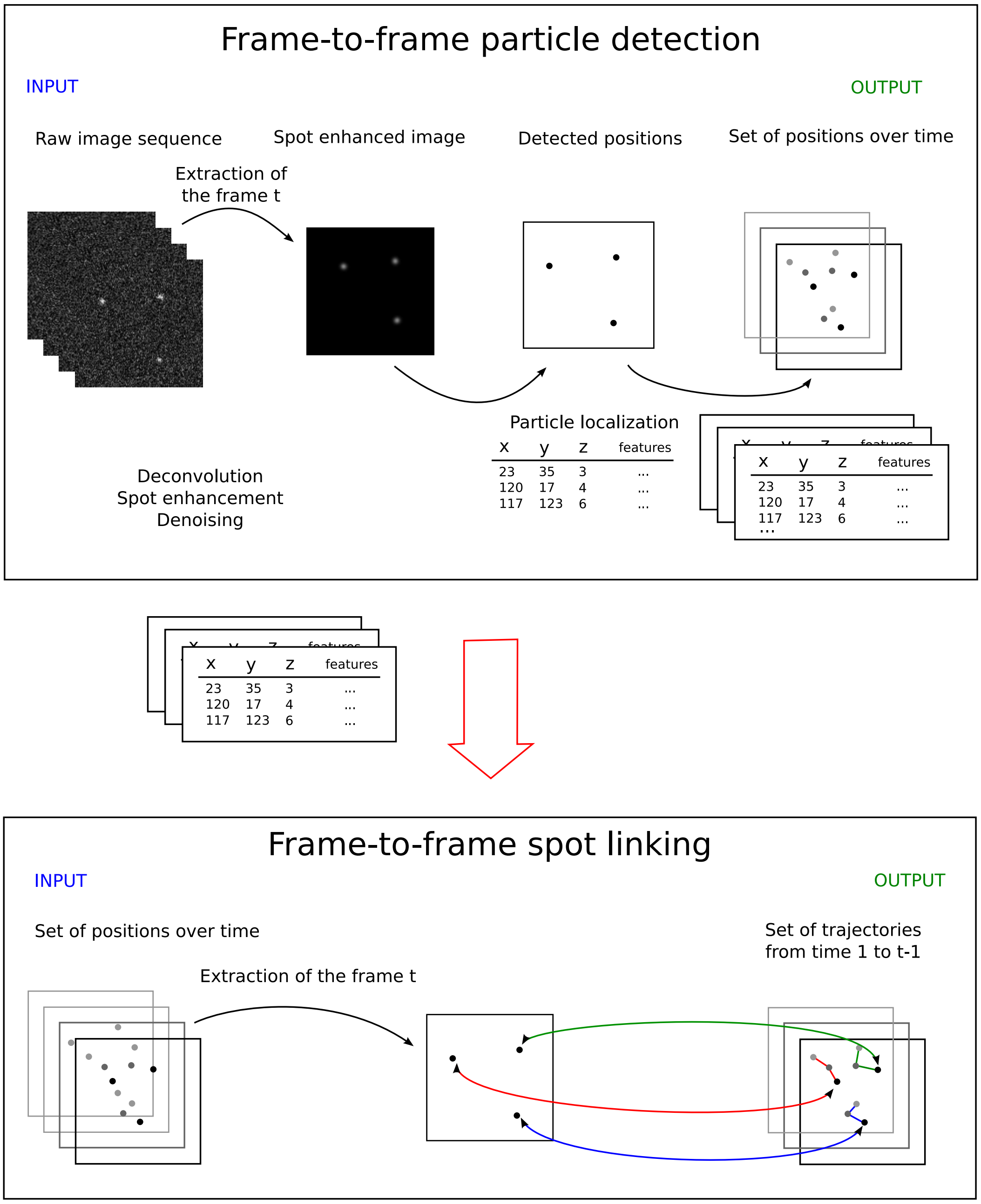


Fig. 1: Standard two-step tracking procedure. Particle locations are first detected in the whole image sequence. Detections are then sequentially linked to form tracks.

Moreover, both the intracellular and extracellular areas are very crowded environments and the fluorescent labeling is often not specific to the targets. Hence, structures such as membranes or filaments may emit photons which create a structured dynamic background in images. In this case the background may obliterate the particle signal, and it may wrongly be considered as a particle. A number of examples showing the diversity of fluorescence images and illustrating the challenges of particle detection are given in Figure 2.

The difficulty of detecting particles generally yields corrupted sets of detections : some particles

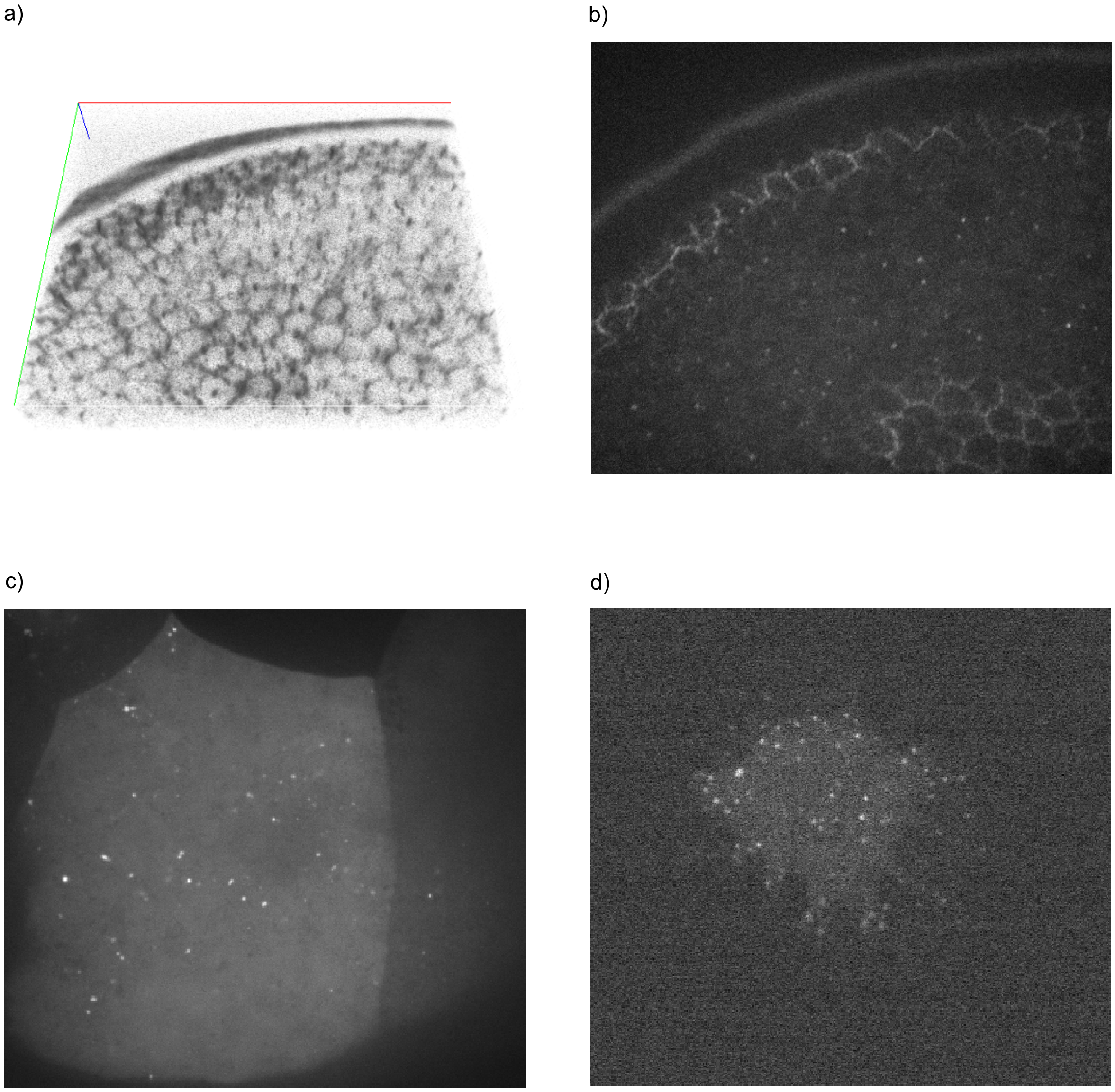


FIG. 2: Examples of fluorescence microscopy images. a) GFP labeled notch protein in the Dorsophila melanogaster oocyte. b) A slice from the 3D stack a). c) *Oskar* mRNA in the Drosophila oocyte. d) Signaling molecule SLP76 fused to the YFP in a lymphocyte.

are missed by the detection procedure and some artifacts are wrongly considered as being some spots of interest, which have to be dealt with. The corruption of the detection sets results in an increased complexity of the frame-to-frame spot linking procedure : the creation of tracks containing false detections should be alleviated, while missing detections have to be compensated for. Second, linking detections into tracks may be ambiguous due to the high density of particles that are often met in biological samples. Finding the links between detections in subsequent frames is generally done by assuming a given model of motion for particles. In practice, modeling the movements of particles is however complicated by the wealth of motion types that nanometric objects can exhibit in cells : free diffusion is often observed, some objects may be confined to a very small volume, while particles may move on a very long distance with a high velocity, generally due to some interactions with the cytoskeleton. Moreover, abrupt transitions between different motion types are commonly observed for a given object.

In practice, the low intensity of nanometric scale objects and the high velocity of their movements result in two opposite requirements : the exposure time has to be long in order to collect enough light from the sample, while the acquisition should be kept as short as possible as to facilitate the tracking of fast particles. A good trade-off is however barely achievable when imaging fast process and a typical experimental data set generally suffers from both poor imaging conditions and a low acquisition rate as compared to the particle movements.

2.2 Recent works

Considerable efforts have been paid in the bioimaging community to answer to the needs of biologists for efficient tracking tools (see [119], [41] and references therein), which have given birth to a wealth of different solutions. In order to overcome the problem of corrupted detections, methods that do not use a frame-to-frame detection correspondence have been proposed to perform simultaneously detection and tracking [143, 25, 136]. The key idea is to represent the image sequence as a spatio-temporal volume in which trajectories appear as smooth curves. This approach is robust to transient object disappearance and poor imaging conditions thanks to a good integration of the temporal information. However, a number of regularity assumptions, such as the smoothness of trajectories in the spatio-temporal volume, are made but are rarely consistent with biological applications. Their use is consequently limited to very few biological cases.

At the same time, iterative two steps tracking methods have witnessed important developments with the use of the Bayesian framework [67, 155], in which a track is modeled as a stochastic sequence of noisy measurements. The main idea is to model further the tracking problem with statistical concepts, whereby estimation techniques can be used to either extend tracks iteratively with the most probable detection [67], or to build iteratively a probabilistic density function of target positions [155] from which the tracks can be deduced. The flexibility of the framework has permitted the development of complex models of target kinetics, allowing the tracking of fast moving particles and changing dynamics targets. The ability to track numerous particles in 3D, even when the acquisition rate is low, has made the Bayesian approach a reference method for 3D particle tracking in time-lapse microscopy sequence [6, 126, 120]. However, the robustness to false and missing detections are still open issues.

3 Contributions of this thesis

Particle tracking is progressively recognized as a reference tool in biology for studying dynamic processes at nanometric scale [145]. However, the relatively small number of studies using particle tracking that have yet been published indicates that existing solutions still fail at satisfying the pressing needs of biologists for automatic and reliable methods capable of analyzing the great variety of biological data. Indeed, a considerable work has been devoted to designing tracking algorithms for specific applications, but no approach has yet succeeded in tracking particles in the wealth of conditions that can be met in molecular biology. The purpose of this thesis is therefore to develop novel mathematical methods that broaden the horizon of particle tracking in biology.

In order to alleviate a specialization to few biological data sets, all along the thesis we have paid a special care to the flexibility of the designed solutions. This is why we have adopted the standard two-step tracking procedure shown in Figure 1 since detecting particles and then linking detections into tracks has proved to be a more flexible approach for particle tracking than performing these two steps simultaneously. In this thesis we propose three novel solutions at both the particle detection and tracking stages in order to widen the range of applicability of particle tracking in biology :

1 - Particle detection in structured backgrounds

When imaging sub-cellular processes the fluorescent labeling is often not specific, hence a structured and dynamic background may be visible since the cell is a very busy environment. Standard particle detection procedures however assume that particles appear as small spots superimposed to flat and slowly varying background, so they are not suited to this frequent case. As a solution to this issue we propose a new particle detection algorithm which deals with the presence of structured backgrounds by separating the background from the particle signal. The proposed approach relies on a morphology based source separation approach which exploits the concept of sparsity of representation [164]. Separating the background and the particles enables the analysis of numerous biological data sets which were previously impossible to process with a sufficient quality for the purpose of particle tracking.

2 - Statistical modeling of the particle tracking problem

The statistical framework for tracking offers a flexible and unified means to select and control the models and assumptions on which the algorithms are based. We have thus adopted this framework and we propose an exhaustive modeling of the problem of tracking particles in microscopy images. We provide a global statistical description of the tracking problem (as done in other communities [9]), as well as statistical models which are specific to particle tracking applications. The framework we define includes the description of particle motion which is based on biophysical models, as well as a statistical modeling of the particle existence (called *perceivability* [110]). The statistical framework we propose can be easily adapted to the wealth of different biological cases by changing the models it includes and adjusting their parameters.

3 - Robust tracking in poor imaging conditions with the E-MHT

In many biological applications the poor imaging conditions result in very diffuse statistical models of the tracking problem, hence making difficult the discrimination of good tracking solutions. A frame-to-frame approach, accounting only for past frames would thus result in non reliable decisions. To compensate for the small amount of information contained in each frame we propose to integrate the information contained in several future frames. To do so, we adapt to biological images the Multiple Hypothesis Tracking algorithm (MHT) which has been previously used for radar tracking and video surveillance applications [138, 47, 18]. We have reformulated the MHT principles in the probabilistic framework for tracking by integrating in the MHT the possibility to account for different types of motion and the target perceivability concept. The original implementation of the MHT is known to be extremely time-consuming when the targets are numerous. Hence, the high number of particles to track in microscopy images has led us to propose a novel and fast implementation of the MHT which allows us to track simultaneously hundreds of particles. The proposed algorithm (which we name E-MHT for *Enhanced*-MHT) is able to decipher very challenging tracking cases, such as tracking numerous particles in extremely noisy 3D image sequences, since it takes full advantage of the statistical modeling of the tracking problem we have proposed.

4 Outline of the dissertation

The outline of this work is as follows. In a first part we review the state-of-the-art methods for tracking particles in microscopy images. In Chapter 1 we first present a number of principles of fluorescence imaging, and we browse through the main methods for detecting particles in fluorescence images. Chapters 2 and 3 are then dedicated to the review of a number of particle tracking methods. We distinguish two main families of approaches : techniques that do not involve a statistical modeling of the tracking problem are described in Chapter 2, and probabilistic methods for particle tracking are presented in Chapter 3.

In the second part of the dissertation we present our contributions to particle tracking methods. In Chapter 4 we present a novel particle detection scheme which is well suited for the processing of fluorescence images containing a structured and dynamic background. In Chapter 5 we describe an exhaustive statistical modeling of the tracking problem, which is then refined in Chapter 6 with a model of particle perceivability in fluorescence images. The modeling of the tracking problem is exploited in Chapter 7 with the proposition of a novel MHT algorithm which is validated and compared to a number of state-of-the-art techniques both on synthetic data sets and real 2D and 3D image sequences.

In the last part of the manuscript we present three different biological studies in which the methods proposed in this thesis have allowed establishing novel biological results. Finally, we conclude this thesis in by summarizing our contributions and proposing several extensions to the proposed advances.

Part I

A review of particle tracking methods in microscopy images

Chapter 1

Detection of fluorescent spots

1.1 Fluorescence image model

Detecting the presence of particles in images is a task which is intricately linked to the acquisition protocol since the properties of the different signals present in images can vary for different imaging modalities and different biological samples. We thus first review the basics of fluorescence imaging, and we then define a number of important concepts which are referred to throughout the dissertation.

1.1.1 Fluorescence imaging principles

The key idea of fluorescence microscopy is to specifically activate fluorophore targeting objects of interest and to collect and measure the emitted light to gain information on these objects presence, number, distribution or localization in the sample. However, most molecules in cells do not have any intrinsic fluorescent property, hence fluorescent staining of objects of interest is a prerequisite for their observation, while unstained objects are expected to emit no (or low) light and thus to be (nearly) invisible. We refer to [179] for a good introduction to fluorescence imaging principles, and we highlight here only few key ideas and major breakthroughs.
Immunostaining is one of the earliest widely used technique for biological sample fluorescent labeling. Its principle is to purify the protein of interest, called antigen, covalently label it with a fluorochrome-labeled antibody which is specific to the antigen, and to reintroduce the complexes in the sample. The purification and reintroduction steps however impose some technical constraints which make immunostaining difficult to use for in vivo imaging.
A major breakthrough has occurred in fluorescence imaging with the development of fluorescent proteins as molecular tags which can be genetically fused to virtually any protein of interest (see [114] for a review). Cloning the green fluorescent protein (GFP) gene from jellyfish species genomes to other organisms [37] has opened the way to new approaches for understanding protein function [171]. Using genetic modifications to incorporate fluorescent proteins indeed avoids the protein purification and reintroduction stages required for immunostaining, hence facilitating in vivo imaging. Moreover, the genetic modifications aiming at labeling a targeted protein can be conserved in specific lineage of the species. The fluorescent labeling process is therefore non invasive and non toxic for the studied organism. Over past decades many improvements have been made to the original wild-type GFP protein and other fluorescent proteins [114] : the GFP expression has been improved and the absorbance peaks have been modified to produce a brighter protein called enhanced GFP (EGFP). A number of modifications have also been done to the GFP and other fluorescent proteins to shift the emission wavelength and obtain a wealth

of different colors. Nowadays dozens of different colors are available which allows distinguishing different proteins of interest and opening new ways to study their interactions : protein colocalization studies are such applications exploiting the variety of colors of fluorescent tags. The cyan fluorescent protein (CFP) and yellow fluorescent protein (YFP) are widespread examples of such fluorescent proteins.

One alternative solution to label mRNA molecules is to use molecular beacons which specifically target the mRNA sequence. Molecular beacons are oligonucleotide hybridization probes which contain an internally quenched fluorophore. When the beacon binds to the targeted nucleic acid sequence it unfolds and the fluorescence is activated. Therefore, a major advantage of molecular beacons over other probes is that the excess of unbound probes is not visible in fluorescence microscopy.

Synthetic fluorescent objects are generally less versatile and brighter than biological objects which are genetically fused to a fluorescent protein. Hence, introducing synthetic fluorescent objects in cells is also a frequently used means of obtaining information on the cellular mechanisms. Such objects can be used as immunofluorescent reagents, tracers and flow cytometry for instance. Fluorescent beads generally consist in polystyrene microspheres which are internally labeled with specific dyes. The wealth of available sizes and colors of beads offers a high degree of flexibility for biological experiments. Past years have also seen the development of semiconductor nano-crystal, coined quantum dots (QDs), as fluorophores for bioimaging [56, 121]. The advantages of QDs over standard fluorophores are broader absorption spectrum and a narrower emission spectrum yielding a brighter fluorescence for the same quantity of illumination. Fluorescent proteins are also subject to the irreversible degradation of the fluorescent molecule (photobleaching) which makes them difficult to observe in long experiments. In contrast, QDs are more stable and can be observed for hours. Coating QDs is required to incorporate them cells : they can be encapsulated in phospholipid micelles for instance [56]. The coating also permits targeting QDs to specific proteins. One important characteristics of most QD crystals is their blinking aspects. Indeed, most commercialized QDs may transiently cease to produce photons in a random manner.

Some proteins in cells are also naturally fluorescent and therefore can produce undesirable light during the observation procedure. This phenomenon is called the autofluorescence of cells. Fluorophore photoactivation is generally achieved by the illumination of the scene by a lamp or a laser beam mounted on the microscope device. Controlling the emitted wavelength allows one to specifically activate a type of fluorophore when several ones are present and to reduce some problems such as the phototoxicity and an excessive sample heating. The light emitted from the fluorophore molecules is then collected by the combination of numerous lenses and measured by a photosensitive receptor such as a charge-coupled device (CCD) or photomultiplier tube to create a digital image of the sample.

1.1.2 Fluorescence microscopes

Many optical systems have been designed to obtain nanometric scale precision images, however we focus on three of them in the rest of the manuscript : the wide-field fluorescence microscope (WFFM), the laser-scanning confocal microscope (LSCM) and the disk-scanning confocal microscope (DSCM), since they are the most frequently used acquisition devices in cell Biology due to the high quality images they produce compared to their relatively low cost. Other devices such as the Structured Illumination Microscope (SIM) [82], the photoactivated localization microscope (PALM) [17], stimulated emission depletion (STED) microscope [83] or stochastic optical reconstruction microscope (STORM) [142] provide higher spatial resolution fluorescence images, but are not yet widespread.

The role of the microscope device is both to activate fluorescent proteins staining the sample, and to collect the resulting photons such as the sensor rests on the focal plane. To do so and to correct aberrations several lenses are generally combined. While the input of the optical system is the photons emitted by the sample, it outputs a 3-dimensional fluorescence intensity map which is recorded plane by plane by the camera.

When the emitted light is incoherent the system is described as being a linear and shift invariant system characterized by a transfer function whose inverse Fourier transform is called the Point Spread Function (PSF). Mathematical models have been established for the diffraction-limited PSFs of WFFM, LSCM and DSCM based on wave optics : they can be expressed by means of Debye's integrals [80, 188]. In practice the output of the system can be seen as the convolution of the input density map with the PSF which thus corresponds to the response of the imaging system to a point source of light (infinitely small object emitting photons). Convolving the input with the PSF is a blurring process of the input which is due to the diffraction phenomenon and some imperfections of the optical system particles.

Most microscopy devices acquire 2D images, hence 3D imaging is generally achieved by acquiring successive thin optical sections with different depth in the sample. The acquisition therefore results in a set of 2D images simply called *stack*. Changing the focal plane is generally achieved by moving the biological sample along the $z-$axis of the microscope which has some important consequences on the image analysis : moving the stage and acquiring multiple images considerably decrease the acquisition rate, the voxel size is also generally different in the vertical direction and in the xy-plane.

1.1.3 Image model

The intensity $I(x, y, z)$ measured in the voxel of Cartesian coordinates (x, y, z) can be accurately modeled as a Mixed-Poisson-Gaussian (MPG) random process defined as :

$$I(x, y, z) = \alpha U(x, y, z) + V(x, y, z), \tag{1.1}$$

with :

$$U(x, y, z) \sim \mathcal{P}(\lambda(x, y, z)), \text{ and } V(x, y, z) \sim \mathcal{N}(\mu(x, y, z), \sigma^2(x, y, z)), \tag{1.2}$$

where $\alpha > 0$ is the overall gain of the system detector. The intensity measure is random and follows a combination of a Poisson distribution $\mathcal{P}(\lambda(x, y, z))$ and a Gaussian distribution $\mathcal{N}(\mu(x, y, z), \sigma^2(x, y, z))$. The randomness is mainly due to the random nature of photon counting and the electronic acquisition noise :

1. The randomness of photon counting is inherent to the process of photon emission : emitting a photon during a small time interval is modeled as a random event occurring with a fixed probability and which is independent of photons already emitted. Photon emission can therefore be seen as a stochastic Poisson process with a spatially varying mean $\lambda_p(x, y, z)$.
2. The electronic acquisition noise occurs at the sensor level during the image acquisition. Many steps are indeed necessary for the conversion of a photon capture event into a digital signal, all of which being subject to some noise since they involve analog signals transmission, conversion and transformation. The most prominent sources of signal corruption are : the dark noise which is the liberation of electrons from the CCD substrate in the absence of photons, the read out noise which comes from the imperfect conversion of charge into a voltage signal and the quantification noise coming from the lossy conversion of the voltage in a Q-bits digital signal. The dark noise can be modeled as a stochastic Poisson process of

mean λ_d. In contrast the read out noise is generally considered as following a normal distribution of mean μ_r and standard deviation σ_r, while the quantification noise is generally neglected when encoding signals with 16 bits.

Relating the above description of noise sources to Equation 1.1, we write $\lambda(x, y, z) = \lambda_p(x, y, z) + \lambda_d$ and $(\mu(x, y, z), \sigma(x, y, z)) = (\mu_r, \sigma_r)$. When a high number of photons is counted (long exposure time, high illumination intensity), $\lambda(x, y, z)$ is increased and the Poisson term can be approximated by a normal distribution. The MPG model hence turns out to be purely Gaussian with $U(x, y, z) \sim \mathcal{N}(\lambda(x, y, z), \sqrt{\lambda(x, y, z)})$ in this case.

The photon counting term is composed of several additive terms when acquiring images of particles : the intensity originating from the in-focus object : u_i, the cell autofluorescence intensity : u_a, and the intensity of out-of-focus objects for the WFFM : u_o, each one being random and following a Poisson distribution of mean λ_i, λ_a and λ_o respectively, such that :

$$\lambda_p(x, y, z) = \lambda_a(x, y, z) + \lambda_o(x, y, z) + \lambda_i(x, y, z) \tag{1.3}$$

The image model presented above shows that only the random intensity u_i mainly depends on the presence of the particles and can be therefore used to identify them. The other terms act as random noise that may corrupt the localization process.

1.1.4 Image quality assessment

As shown in Section 1.1.3, the intensity originating from the particles of interest is corrupted by many sources of noise which may be a source of failure for the detection process. It is thus essential to define tools to assess the quality of images and understand the parameters influencing the quality. The signal to noise ratio (SNR) is a standard tool to quantify the quality of a noisy signal : its principle is to compare the amplitude of the signal of interest to the magnitude of the noise. In our case we define the SNR as the ratio of the expected intensity originating from the particles to the standard deviation of the intensity in images :

$$SNR(x, y, z) = \frac{E[u_i(x, y, z)]}{\sqrt{E[(I(x, y, z) - E[I(x, y, z)])^2]}}, \tag{1.4}$$

which can be re-written as :

$$SNR(x, y, z) = \frac{\alpha\lambda_i(x, y, z)}{\sqrt{\alpha(\lambda_i(x, y, z) + \lambda_a(x, y, z) + \lambda_o(x, y, z) + \lambda_d(x, y, z)) + \sigma_r^2}}. \tag{1.5}$$

The SNR expression (Eq. 1.5) shows that if sources of noise are negligible compared to the photons emission of the particles $(\alpha\lambda_i(x, y, z) \gg \alpha(\lambda_a(x, y, z) + \lambda_o(x, y, z) + \lambda_d(x, y, z)) + \sigma_r^2)$ then the SNR scales with the square root of the gain and $\lambda_i(x, y, z)$ since in this case : $SNR(x, y, z) \approx \sqrt{\alpha\lambda_i(x, y, z)}$. The second extreme case is when the mean particle photon emissions $\lambda_i(x, y, z)$ is negligible compared to the other sources of noise $(\alpha\lambda_i(x, y, z) \ll \alpha(\lambda_a(x, y, z) + \lambda_o(x, y, z) + \lambda_d(x, y, z)) + \sigma_r^2)$. In this case the SNR scales linearly with $\lambda_i(x, y, z)$. The true behavior of the SNR is somewhere in between these two cases, as shown in Figure 1.1.

Increasing the SNR can be achieved by increasing the product $\alpha\lambda_p$ by different means : increasing the acquisition gain α and increasing the exposure time or the power of the illumination source. A long exposure time and an increased illumination intensity however result in major issues in the context of particle tracking :

- A high level of illumination light may damage the biological sample, hence invalidating the observation.

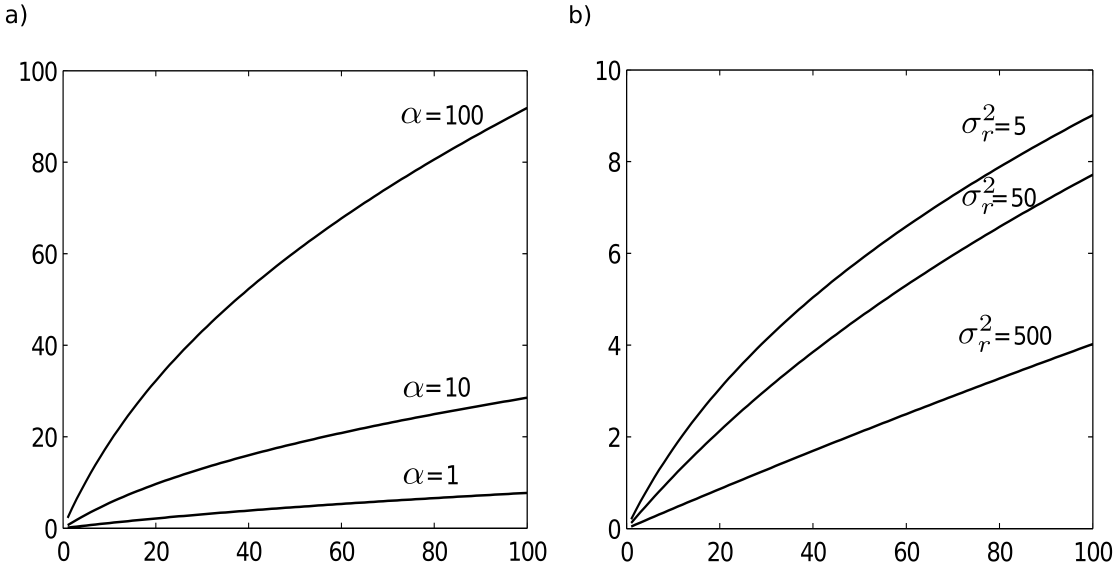


FIG. 1.1: SNR v.s λ_i curves with the image settings : $\lambda_o = 5$, $\lambda_a = 10$ and $\lambda_d = 3$. x-axis : SNR computed with Equation 1.5. y-axis : mean target intensity λ_i. a) SNR for system gain settings : $\alpha_r^2 = 1, 10$ and 100. The read-out variance is set to 50. b) SNR for three read-out variance settings : $\sigma_r^2 = 5, 50$ and 500. The gain α is set to 1.

- An excess of excitation causes fluorophore molecules to degrade and the fluorescent signal to fade over time. This process, called photobleaching, prevents the acquisition of a long image sequence when illuminating samples with a high quantity of light.
- A long exposure time significantly slows down the acquisition process. For 3D particle tracking applications this issue is critical since particle movements may be very fast compared to the acquisition rate. As we demonstrate in Chapter 2 tracking performance is intricately related to the acquisition rate, a poor temporal sampling would thus result in an increased difficulty of the tracking problem.

Setting up the acquisition device is therefore a matter of trade-off between a high image quality to facilitate the detection process, a low level of illumination to preserve the biological sample and its fluorescence capability, and a short exposure time to follow fast moving particles.

In practice parameters characterizing the acquisition process such as λ_d, μ_r and σ_r could be estimated beforehand with some experiments, but are in practice rarely available. Also λ_o and λ_a are sample dependent and difficult to estimate. For a particle at a given time and position we get only one realization of its photons emission random process, hence making the accurate estimation of $\lambda_i(x, y, z)$ impossible. For the study of the accurate estimation of MPG systems we refer to [170] and [186]. The difficulty of estimating $\lambda_i(x, y, z)$ and the other MPG process parameters leads to the definition of a second SNR criterion for particle detection applications, called peak SNR (PSNR), which is easy to estimate in acquired images without any specific knowledge on the acquisition device. The PSNR of a particle at position (x, y, z) is defined as :

$$PSNR(x, y, z) = \frac{I(x, y, z)}{\sigma_n}, \tag{1.6}$$

where σ_n is an empirical estimation of the noise standard deviation : $\sqrt{\alpha(\lambda_a + \lambda_o + \lambda_d) + \sigma_r^2}$, obtained by the estimation of the MPG process parameters in an empty and flat neighborhood of the particle.

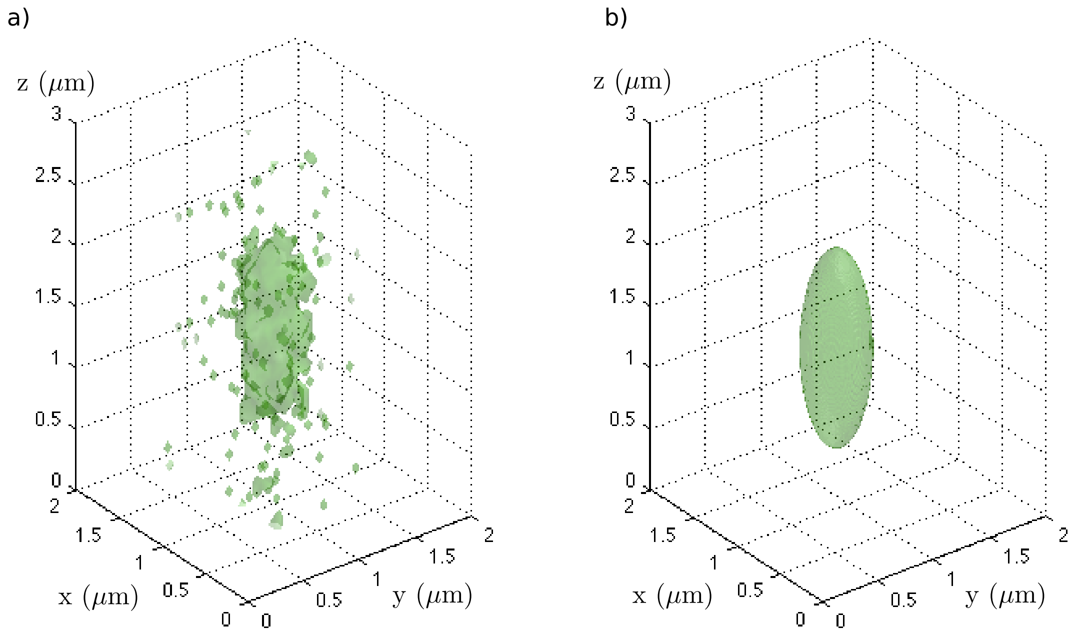


FIG. 1.2: 3D PSF of the DSCM and its Gaussian approximations. a) The observed microscope PSF with a 100x lens. b) The fitted 3D Gaussian function.

1.1.5 Simplified additive and Gaussian image model

In the context of particle tracking the observed objects are small compared to the acquisition device resolution and are nearly isotropic. Particles however appear as blurred spots instead of well defined points due to the diffraction phenomenon and some imperfections of the optical system particles which determine the microscope PSF. In most cases, for WFFM, LSCM and LSCM the theoretical PSF can be accurately approximated by a multidimensional Gaussian function which offers some important computational facilities [188]. Thus, the following simplified model can be used in place of the MPG model (Eq. 1.1) :

$$I(x,y,z) = T(x,y,z) + W(x,y,z). \tag{1.7}$$

The signal of objects of interest at position (x,y,z), $T(x,y,z)$, is considered as a deterministic value computed as :

$$T(x,y,z) = \sum_{i=1..n} P_i(x,y,z), \tag{1.8}$$

where $P_i(x,y,z)$ is the signal originating from the i^{th} particle in pixel (x,y,z). When a particle is significantly smaller than the resolution of the microscope, its intensity profile P_i is well represented by the Gaussian PSF of the microscope with a specific amplitude A_i :

$$P_i(x,y,z) = A_i e^{-\frac{(x-x_i^0)^2-(y-y_i^0)^2}{2\sigma_{xy}^2}-\frac{(z-z_i^0)^2}{2\sigma_z^2}}, \tag{1.9}$$

where (x_i^0, y_i^0, z_i^0) is the coordinate of the i^{th} particle. As an example, in Figure 1.2 the experimental PSF of a DSCM is given, and we show that the observed function is well fitted by a 3D Gaussian function. The resolution of the microscope is generally coarser along the z-axis so the

profile is isotropic in the xy-plane but elongated in the z direction ($\sigma_z > \sigma_{xy}$).
In the additive Gaussian model (Eq. 1.7), $W(x, y, z)$ is a white and random term following a normal distribution $\mathcal{N}(\mu_n, \sigma_n)$. Hence, the distribution of $W(x, y, z)$ should summarize the different sources of random variations described by the MPG model. To do so the moments μ_n and σ_n are generally empirically estimated in the images.

1.2 Detection methods

As shown by the MPG image model (Eq. 1.1) and the simplified Gaussian model (Eq. 1.7) the existence of a fluorescent particle is reflected by a local increase of the intensity in the image. Due to the low particle signal and the high level of noise (low SNR conditions) this intensity increase may be difficult to detect pixel wise, so pixel wise detection methods generally lack reliability. As an example, thresholding independently each pixel intensity value to obtain a binary mask of the target presence would not provide satisfactory results with most noisy images. A number of much advanced detection methods have thus been proposed to identify the location of spots in fluorescence images (see [156] and references therein for a comparative review). We highlight in this section few of them with a particular methodological interest.

1.2.1 Optimal filter design for known particle profile

The signal of particles in convoluted by the PSF during the acquisition stage, hence particles appear as extended blurry spots in images instead of compact and bright dots. Increasing the amount of information at a given position is therefore possible by accounting for intensities of neighboring pixels : they should be bright if a target is present at the considered position. Filtering methods exploit this idea by integrating out intensities following a specific kernel function. For instance in [143] an optimal filtering based detector is presented for the simplified Gaussian model (Eq. 1.7). The principle is to design a filter that best correlates with spots shape to obtain the maximum PSNR.
In [143], the authors have experimentally found that the spectral density of fluorescence microscopy images is isotropic and well represented by a power law $S_n(\omega) \propto ||\omega||^{-s}$ with the fractal exponent $s = 1.8$ (where ω is the radial spatial frequency). This spectral density is indicative of a colored noise W in Equation 1.7. It is thus proposed to apply first a pre-filter $H(\vec{\omega}) = ||\vec{\omega}||^{s/2}$ to transform the $1/\omega^s$ noise W in an uncorrelated noise. Once the noise is whitened a standard correlation detector is applied to enhance spots corresponding to the particles of interest. If P is the intensity profile of a particle then this matched filter is the complex conjugate of the Fourier transform of P : $||\vec{\omega}||^{s/2}\hat{P}^*(\vec{\omega})$.
It is shown in [143] that the whitening filter and the matched filter can be combined in a single filter, coined whitened matched filter, whose response is proportional to $\Delta^{s/2}P(x, y, z)$ where $\Delta^{s/2}$ is the $s/2$-fractional iterate of the Laplacian operator. With the approximation $s = 2$ and using the Gaussian definition of a particle profile (Eq. 1.9) the optimal detector corresponds to the Laplacian of a Gaussian (LoG), or 'Mexican hat' filter, which formula in the space domain is the following :

$$\mathrm{LoG} = \Delta P(x, y, z) = \frac{1}{\sqrt{(2\pi)^3\sigma_{xy}^4\sigma_z^2}}\left(\frac{x^2+y^2}{\sigma_{xy}^4} - \frac{2}{\sigma_{xy}^2} + \frac{z^2}{\sigma_z^4} - \frac{1}{\sigma_z^2}\right)e^{-\frac{x^2+y^2}{2\sigma_{xy}^2} - \frac{z^2}{2\sigma_z^2}}. \tag{1.10}$$

Figure 1.3 shows a $2D$ LoG filter for a Gaussian PSF standard deviation $\sigma_{xy} = 1.5$ pixels. The convolution kernel is highly peaked, hence enhancing high frequency signals, and isotropic with a

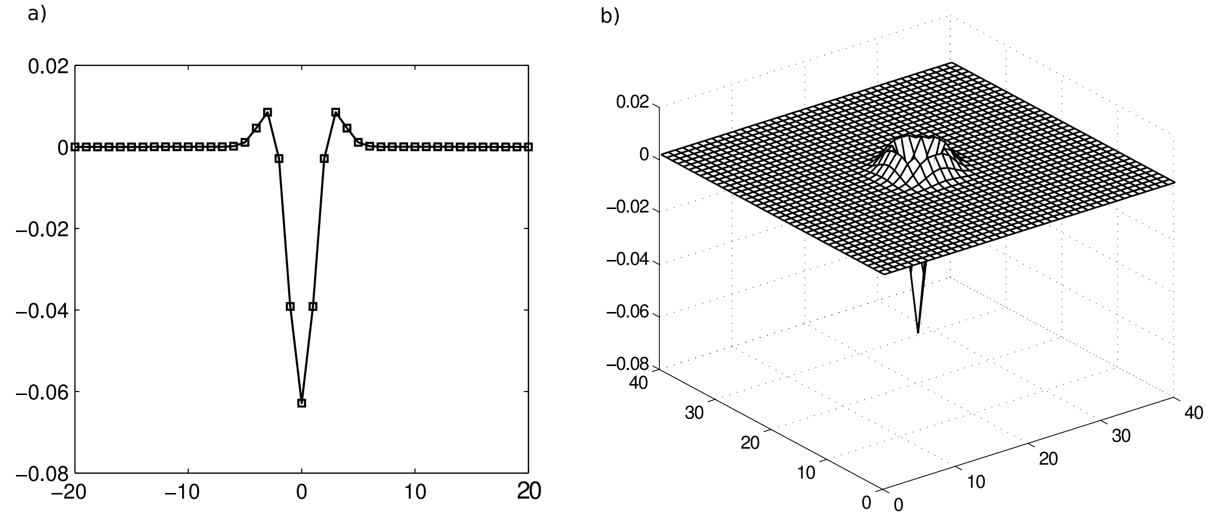


FIG. 1.3: 2D Laplacian of a Gaussian filter with $\sigma_{xy} = 1.5$. a) Section of the filter along the x-axis. b) 3D view of the 2D LoG kernel function.

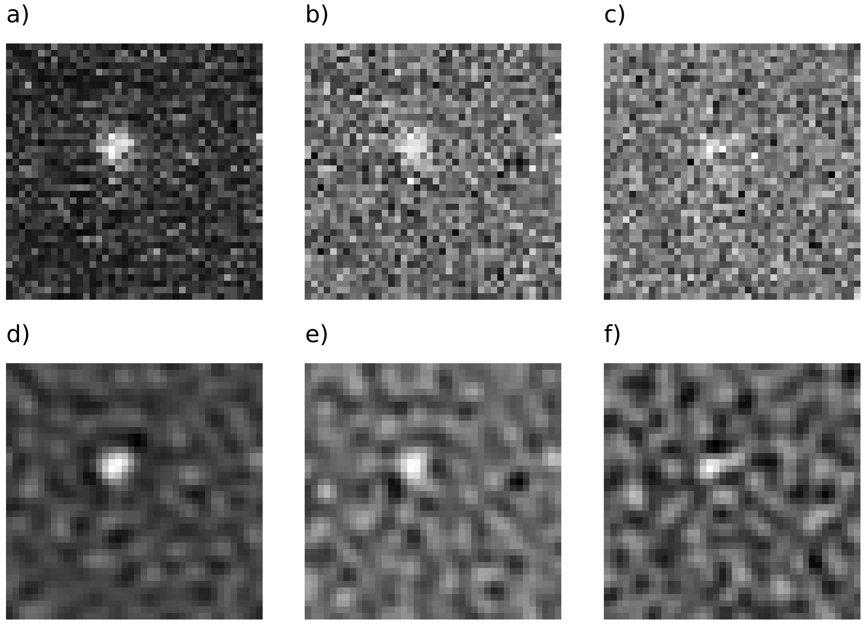


FIG. 1.4: LoG filtering of a fluorescent confocal image of bead. a) Confocal image of a fluorescent bead. d) The convolution of the image a) by the 2D LoG filter displayed in Figure 1.3. b)-c) images resulting from the addition of a synthetic Gaussian noise (of standard deviation 200 and 300 respectively) to the image a). d)-e) LoG filtering of images b) and c) respectively.

null sum whereby slowly varying backgrounds are flattened. In practice the $3D$ LoG filtering can be achieved by successive 1D convolutions along the x-axis, y-axis and z-axis with a 1D kernel which has a similar shape as the kernel section displayed in Figure 1.3 a). The separability of the 3D convolution makes it fast to compute even when processing large images.

We illustrate the spot enhancement properties of the LoG filtering in Figure 1.4 a) and d) : a confocal image of fluorescent bead is convolved with the LoG filter displayed in Figure 1.3. In the whitened matched filtered image the spotty signal is enhanced while the noise is flattened. The experience is repeated with the same bead image in which a synthetic Gaussian noise has been added to decrease the SNR (Figure 1.4 b) and c)). Convolving noisy images with the LoG filter still enhanced the spot despite the high level of noise as illustrated in Figure 1.4 e) and f). Spots can thus be reliably detected in the filtered image, for instance by selecting local maxima exceeding a predefined threshold. A number of methods also rely on extracting some features in spot enhanced images and then using a supervised [154] or non supervised [173] classification method to split local maxima in two classes : noise and particle induced maxima.

1.2.2 Separation of overlapping spots by fitting a mixture model to the image

When two particles are very close to each other the superposition of their blurry profiles may give rise to a single local maximum in the spot enhancing filtered image. The idea introduced in [173] is to build a candidate set of mixture models to resolve closely spaced particles. The mixture definition is based on an additive Gaussian image model : a mixture model consists of a superposition of n shifted Gaussian functions with the same shape as the microscope PSF. The mixture intensity at position (x, y, z) is given by :

$$M(x,y,z) = \sum_{i=1..n} P_i(x,y,z) + b, \tag{1.11}$$

where b is an estimated constant value modeling the background. In practice the number of Gaussian profiles and the position of their centers have to be estimated to fit the observed image. A trade-off has however to be found between the goodness of fit of the model and the complexity of the mixture to avoid the addition of non significant models fitting the noise. To do so a bottom-up approach is used in [173] : new candidate profiles are sequentially added until the improvement of the fit is no longer significant.

For M_n, a model of order n, its $\chi^2_{r_n}$ statistics is computed as the sum of squared error divided by the degree of freedom r_n :

$$\chi^2_{r_n} = \frac{\sum_{(x,y,z)\in I}(M_n(x,y,z) - I(x,y,z))^2}{r_n} \text{ with } r_n = N - p_n, \tag{1.12}$$

where N is the number of voxels and $p_n = 4n+1$ denotes the number of parameters of the model. A good model M_n should result in a low residual $\chi^2_{r_n}$ statistics. To decide whether a model M_{n+1} fits significantly better the image I a statistical test is formulated to test the equality of their χ^2 statistics :

$$\begin{aligned} H_0 &: \ \chi^2_{r_n} = \chi^2_{r_{n+1}} \\ H_1 &: \ \chi^2_{r_n} > \chi^2_{r_{n+1}} \\ T &= \frac{\chi^2_{r_n}}{\chi^2_{r_{n+1}}} \propto F \end{aligned} \tag{1.13}$$

The test statistics T is Fisher distributed and the model addition procedure is stopped in practice as soon as T falls beyond the confidence level of 95%.

By using a Gaussian profile the particle location is estimated with a much higher accuracy than approximating it as being a local maximum in a filtered image or as the center of mass of the neighboring pixels. Indeed when the SNR is high, the precision of the localization procedure reaches the single nanometer for isolated particles [173]. Using the PSF information also allows one to split spots with a distorted shape in several closely spaced detections : in high SNR conditions the capability of the method to resolve particles separated by a distance which is a factor 2 below the lateral Rayleigh limit in the $xy-$plane is demonstrated in [173]. In confocal microscopy the precision of the detection algorithm is as low as 5 nanometers in high SNR conditions.

1.2.3 Multiscale image analysis

The wavelet transform for multiscale image representation

We show in Figure 1.5 a) an DSCM image of *Drosophila melanogaster* oocyte in which Golgi units have been genetically labeled with GFP. Golgi units appear as bright spots in images, however, the labeling has also integrated membranes so the delineation between the oocyte and nurse cells produces a large amount of light which results in a non uniform background to its vicinity. As shown in Figure 1.5 a) some Golgi units are well separated and thus appear as small isotropic spots. However, near the canal rim separating cells they often tend to aggregate and form much larger spots which cannot be approximated by the microscope PSF anymore. The detection methods presented above are thus not suited to this case since they rely on pre-defined and fixed particle template : the PSF. The issue of detecting particles with various shapes and sizes in a non-uniform background is often encountered in biology, hence detection methods which are robust to particles size and shape variations are highly desirable. Methods based on the wavelet analysis (see [117] and references therein) have been designed to fulfill this multiscale requirement.

Wavelet filtering is a multiscale analysis method : the image is filtered by a kernel function which is dilated several times whereby signals with the same shape but different sizes are enhanced in different wavelet scales. The wavelet decomposition results in a representation of the image in a specific mathematical space from which an inverse transform exists, allowing one to exactly reconstruct the original image from the wavelet coefficients. The whole image information is contained in the wavelet representation and we thus say that the wavelet transform is a representation dictionary, or simply a dictionary, of the image domain.

The *à trou* algorithm for isotropic and overcomplete wavelet decomposition

The *à trou algorithm* [85, 153] is an efficient scheme for computing coefficients of discrete and redundant wavelet decompositions. The principle of the *à trou algorithm* procedure is to compute wavelet coefficients at scale j, also coined detail image d_j, as the difference of two low pass filtered images a_j and a_{j-1} called scale images. An iterative scheme is applied as follows :

$$\begin{cases} a_j = \bar{h}^{\uparrow j-1} \star a_{j-1} \\ d_j = a_{j-1} - a_j \end{cases}, \tag{1.14}$$

with j varying from 1 to the desired number of scales J. The function h is a low-pass filter, $h^{\uparrow j}[l] = h[l]$ if $l/2^j \in \mathbb{Z}$ and 0 otherwise, $\bar{h}[n] = h[-n]$ and "$\star$" denotes the convolution. The backward wavelet transform is simply in this case the sum of the J detail images with the last scale image :

$$I = a_J + \sum_{j=1..J} d_j, \tag{1.15}$$

which is an extremely cheap procedure. The forward transform itself is also very efficient : the wavelet is separable, hence leading to the computation of the j^{th} scale image by a succession of

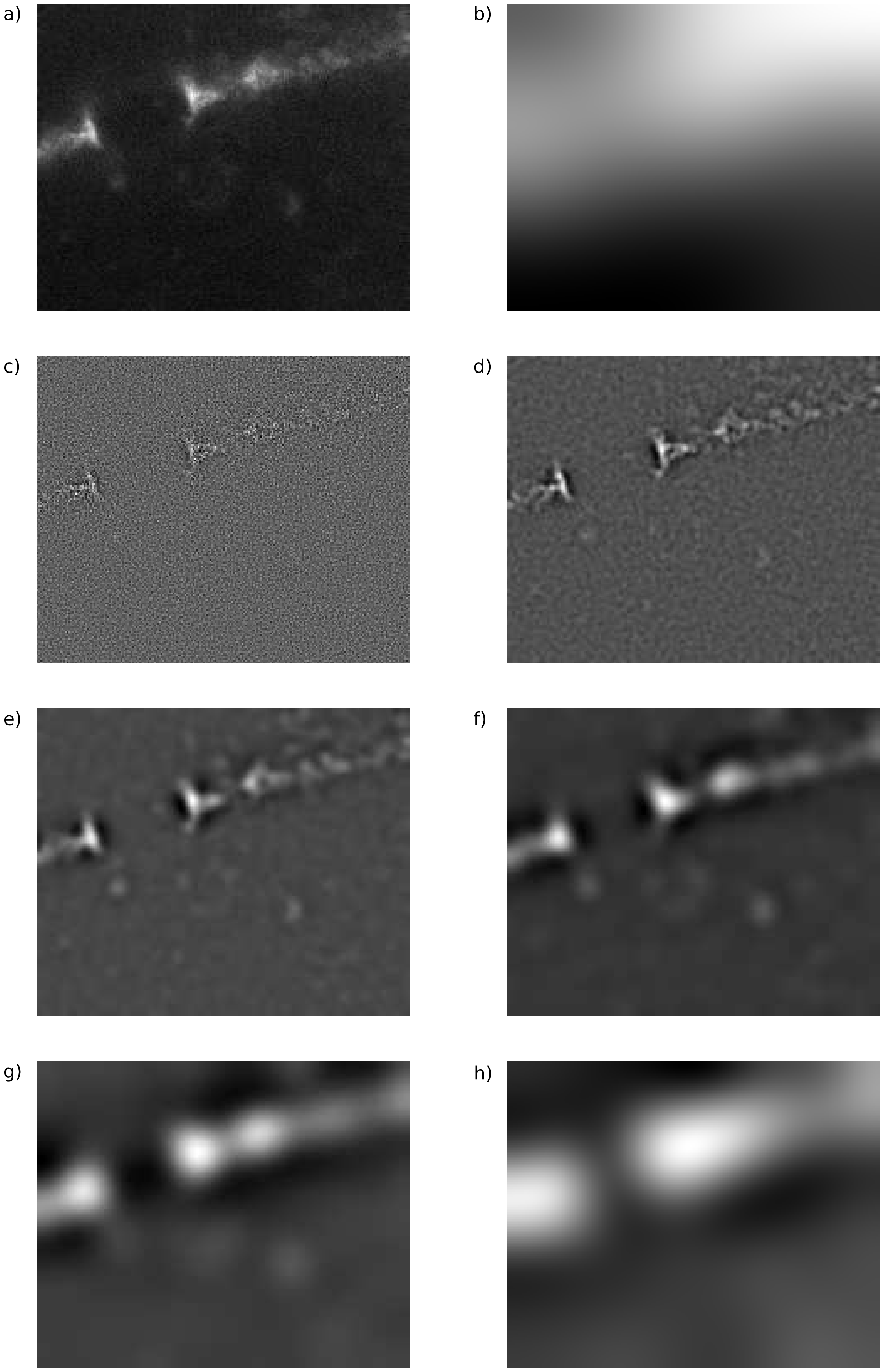


FIG. 1.5: Image decomposition using the B3-UDWT. a) confocal spinning-disk image of a Drosophila oocyte with GFP-labeled Golgi units. b) Approximation image of the B3-UDWT with six scales of a). c)-h) Detail images of the B3-UDWT of a) from scale 1 to 6 respectively.

1D convolutions with the 1D filter $\bar{h}^{\uparrow j-1}$ along each direction.
When using the *à trou* decomposition the wavelet transform is said *undecimated* with comparison to the orthogonal decomposition scheme which is *decimated* [117]. We thus denote UDWT the resulting undecimated discrete wavelet transform. Using a non decimated scheme results in a redundant representation which has some interesting properties for image denoising or target detection applications. For instance, one major advantage is shift-invariance of the representation.

B3-spline wavelet for fluorescent spots representation

Representing particles with few high amplitude coefficients is of prime importance when using a wavelet decomposition of images for particle detection applications. Indeed, if the signal of one particle is represented with a large number of low value wavelet coefficients, discriminating the particle contribution from the noise and the background would be hardly feasible. The choice of the wavelet function can therefore have a dramatic impact on the image analysis : the wavelet function has to fit well to particles shape in order to represent particles with high amplitude coefficients which are easily identifiable. For fluorescence microscopy images processing the B3-spline based overcomplete wavelet [169, 170] fulfills this requirement : this isotropic wavelet transform is indeed characterized by a wavelet function which is very close to the negative LoG filter, as shown in Figure 1.6, and therefore correlates well with the microscope PSF. The wavelet decomposition corresponds in this case to a filtering of the images with a LoG filter dilated to different extents. Each wavelet scale enhances the isotropic features of a given size range. As an example, in Figure 1.5 the B3-UDWT of the *Drosophila* oocyte image is illustrated. The multiscale decomposition property of the UDWT is highlighted : small vesicles are enhanced in scales 2 and 3, while big aggregates produce high coefficients in scales 3 and 4. The approximation image only contains the slowly varying background. Detail images 2, 3 and 4 therefore contain the relevant information for Golgi unit localization and should be analyzed during the detection process.
The B3 wavelet is moreover capable of canceling flat and linear extended signals : the background does not influence the wavelet coefficients but only a residual called *approximation image* which is not used in the analysis. This property comes from the two first moments of the wavelet function which are null [168]. Background cancellation is a key property of the B-spline based wavelets for fluorescence image processing because the wavelet scales are invariant to very low frequency signals which are often observed in this imaging modality. Indeed the illumination is rarely perfectly uniform, the background value may change between images, and some very big and bright objects may produce a non uniform background to their vicinity. We show for instance in Figure 1.5 that the membrane between cells produces a high level of intensity to its vicinity, especially in the ring-canal, but this slowly varying background is flattened in the small wavelet scales. The representation of particles is therefore not influenced by the proximity of this important source of light.

Detecting spots by multiscale wavelet products

The overcomplete wavelet transform results in a redundant representation of the image : the number of coefficients is several times above the size of the image and coefficients are mutually correlated. The idea presented in [129] is to exploit the interscale correlation of the coefficients obtained with the overcomplete $B3$-UDWT to identify spots in a robust manner. Indeed, the white Gaussian noise contained in the image does not propagate through scales while the signal of interest does [98]. In [129] it is thus proposed to select multiple wavelet scales corresponding

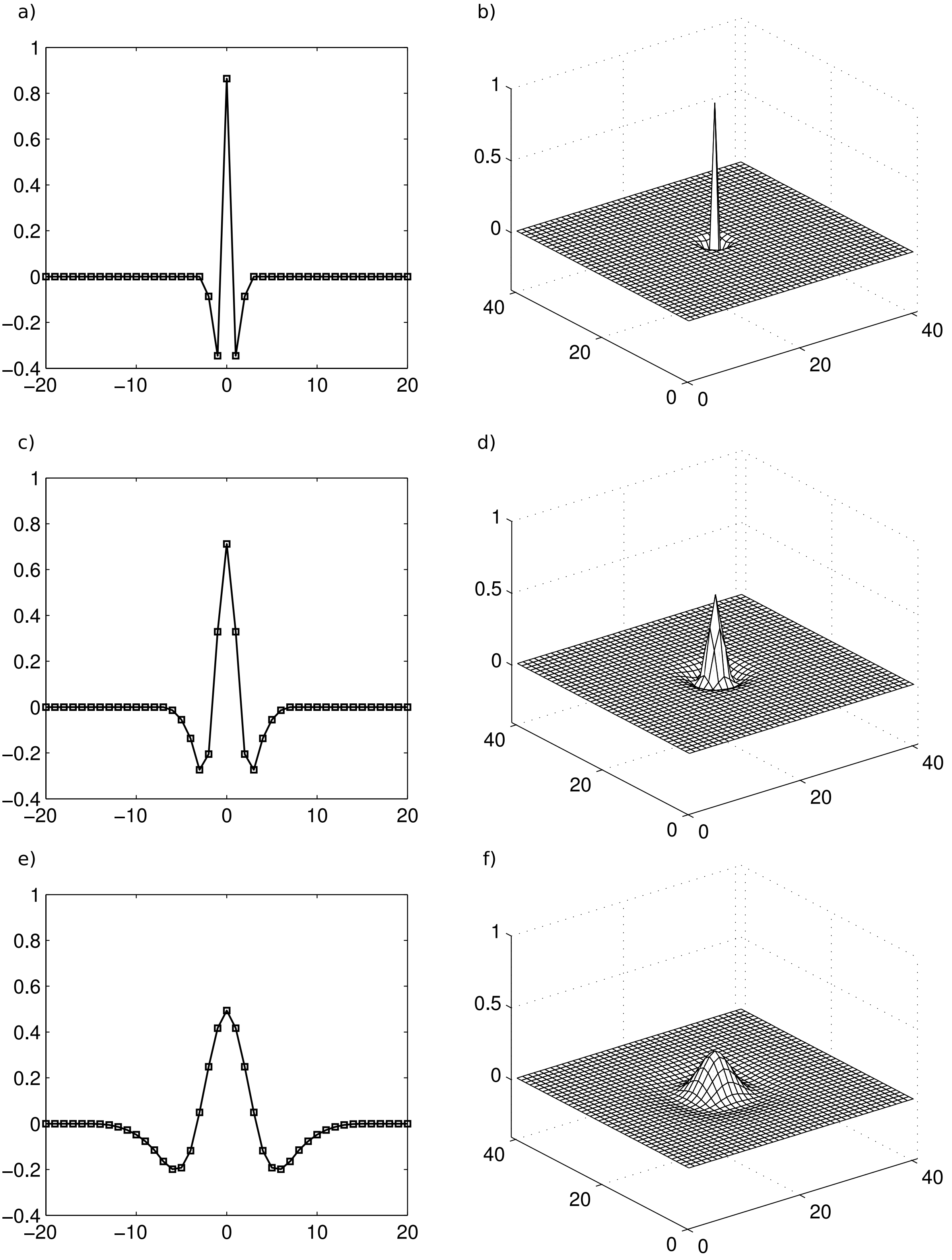


Fig. 1.6: B3-spline isotropic wavelet functions. a), b) and c) : wavelet function in 1D for the scale 1, 2, and 3 respectively. b), d) and f) : corresponding wavelet functions in 2D. In 1D and 2D the wavelet function is nearly a dilated version of the LoG filter.

to the range of particle sizes, denoise them and compute the inter scale product. This results in high coefficients at particle locations only which are then detected by means of a threshold. The good performance of the algorithm has been demonstrated in several types of immunomicroscopy images.

MPG image denoising in wavelet scales

When exposure time is short and the excitation light is low, the Poisson noise in the MPG image model cannot be approximated accurately by a Gaussian term since λ_p is low. Unfortunately most detection algorithms are dedicated to the Gaussian case, hence leading to poor performances in the Poisson case. A solution to this issue is proposed in [186], it relies on a wavelet analysis technique combined with a procedure, coined Variance Stabilizing Transform (VST), to transform the MPG process into a near Gaussian process with a stabilized variance. The principle of the detection method is to apply the undecimated $B3$-UDWT on the image containing Poisson noise, denoise the coefficients and reconstruct a noise free image which is then analyzed for particle localization.

When applying the standard wavelet transform to a MPG process the resulting coefficients are still a MPG process which is difficult to analyze. This issue is solved in [186] by modifying the standard *à trou algorithm* such that a VST is applied to the scale images when computing wavelet coefficients. The modified *à trou algorithm* is the following :

$$\begin{cases} a_j = \bar{h}^{\uparrow j-1} \star a_{j-1} \\ d_j = T_{j-1}(a_{j-1}) - T_j(a_j) \end{cases}, \tag{1.16}$$

where T_j is the VST operator for the low pass image at scale j. Applying the VST the approximation image a_j yields a nearly pure Gaussian noise distribution in the band $T_j(a_j)$, and it is shown in [186] that the Gaussianity is transferred to the detail image. The Gaussian distribution of the noise in the detail image d_j allows one to apply a simple and efficient denoising statistical operator to it. For instance, in [186] a multiple hypothesis testing approach based on the control of the false discovery rate [16] is proposed to detect statistically significant coefficients and to set to 0 non significant ones. The denoised detail images are then combined to recover a noise free image in which spot positions are easily identified. The benefits of the approach have been demonstrated with the detection of small endocytic vesicles in WFFM images.

1.2.4 Other detection methods

We have highlighted above a few detection techniques which are closely related to the characteristics of the acquisition device and have optical and statistical foundations. Less specific spot detection techniques also exist as listed in [156] which compares the most used and effective spot detection techniques. The distinction is made between unsupervised spot enhancing and supervised spot detection techniques. On top of LoG filtering, wavelet product detection and VST based denoising, unsupervised techniques also include :

- (pseudo-)Top-Hat filtering [29, 30] : the filter discriminates the spots relying on their round shape by comparing the mean intensity into a disk of radius r_{top} to the mean intensity of the circular region of radius range $r \in]r_{top}, r_{brim}[$. The difference between the two mean intensities is thresholded to obtain a binary segmentation of the image.
- Grayscale opening Top-Hat filtering : after a Gaussian smoothing the grayscale opening [149, 161] of the image is done with a flat disk of fixed radius r_A. The top hat filtered image is then obtained by subtracting the grayscale opening to the smoothed original image. Only

compact structures which are smaller than the structuring disk remain, hence discarding large background elements.
- H-Dome based detection [157, 160] : the use of the H-Dome transformation [178] results in an image where all objects have the same maximum intensity. In practice a LoG filter is first applied to enhance particle signals and flatten the local background. The grayscale reconstruction [178] is then applied on the filtered image to keep only small blobs around local maxima of height superior to the threshold h. In [157] and [160] the filtered image is used as a probability map of particle presence which is used by the probabilistic particle tracking algorithm to sample positions. Since one spot results in the sampling of several positions the mean-shift algorithm [45] is then used to cluster them and recover single particle positions.
- Local curvature analysis : due to the roundness of spots the local intensity curvature can be combined with the intensity level to build a map of particles position [113, 173, 155]. The curvature, given by the determinant of the intensity Hessian matrix of a local window, is low for random distributions of intensity changes which are typical of noise induced local maxima.

Spot enhancing techniques described above have common characteristics : spot enhancement is based on prior shapes of the spots and of the background, which are exploited by different means. Discriminating between spots originating from particles signal and a random local intensity increase due to the noise is then achieved by fixing some sort of threshold on the enhanced signal. In contrast a class of supervised algorithms does not use such prior knowledge to classify data : the Adaboost algorithm [62] is based instead on a large number of features, called weak classifiers, which are adaptively selected or discarded to create the enhancement function and find the optimal classifying function. The lack of *a priori* description is compensated for by the intervention of the user who has to label a training set of images. Supervised spot detection with Adaboost is proposed in [96] and the Fisher discriminant analysis [118] based classification is used instead in [154].

1.3 Summary

Fluorescence images are random by nature : the randomness is intrinsic to photons emission and the acquisition device also corrupts the signal in a random way. When the level of randomness is low compared to the emitted signal (high SNR) most spot detection methods produce satisfactory results [156] since the image is very reliable. Refined techniques exploiting strong prior information on the signal shape, such as mixture fitting for overlapping spots separation, can be used. However, when imaging fast biological processes these conditions are rarely met : particles have low intensity due to the fast acquisition and the presence of other objects results in a high intensity background. In such low SNR conditions the image information is no longer reliable and accounting for a wealth of prior information is required. Indeed, information such as the type of noise (Poisson and/or Gaussian), the shape of the targets (isotropic or elongated) or their size will play a key role in the discrimination between targets and noise. Supervised detection is an example of the tailoring of the algorithms to the image characteristics : the detection procedure is specifically fitted to every single data set.

Signal randomness imposes some inherent limitations to detection procedures. Indeed in very low SNR condition the noise level is so high compared to the signal of interest that this one may be irremediably lost. For instance electronic noise may randomly obliterate low intensity particles. Spot-like artifacts can also be randomly produced. In such cases detecting artifacts and missing some particles cannot be avoided, whatever the used detection procedure. Tracking algorithms must therefore be able to accommodate numerous false and missed detections when processing poor quality images.

Chapter 2

Non-probabilistic particle tracking in microscopy images

2.1 Tracking particles as solving a point correspondence problem

An intuitive idea to track a set of particles is to detect the targets in the whole sequence of length l with one of the algorithms presented in Chapter 1, and then to link these detections into tracks. At each frame one has to associate the set of detections to the active tracks, so tracking is often viewed as a standard point correspondence problem. A track is defined here as a succession of detections through time. As an example, we note $z_{t_i}^k$ the ith track which is composed of the set of k detections from time 1 to time k :

$$z_{t_i}^k = \{z_{t_i}(1), z_{t_i}(2) \dots z_{t_i}(k)\}. \tag{2.1}$$

We note Z^l the set of detections in the whole sequence : $Z^l = \cup_{k=1..l} Z(k)$, where $Z(k)$ is the set of detections in the frame k.

2.1.1 Automatic tracking by nearest neighbor linking

We first suppose that the detection algorithm performs ideally : the set Z^l contains a detection from each target in each frame, with no false detection. In this case, and when the acquisition rate is fast compared to the targets dynamics, we expect a target to move on a small distance, so the next target detection should be at very close place to the location of the target in the previous frame. An intuitive idea is thus to extend iteratively each track with the detection that is the closest to the previous track detection. For instance, $z_{t_i}^k$ is extended at time $k+1$ with the detection in $Z(k+1)$ with the minimal Euclidian distance to $z_{t_i}(k)$. When tracking n targets, the minimal distance is appropriate only if the distance between two consecutive detections of the same target is smaller than the distance to other targets :

$$\forall j = 1..n \ s.t \ j \neq i : ||\mathbf{z}_{t_i}(k+1) - \mathbf{z}_{t_i}(k)||_2 < ||\mathbf{z}_{t_j}(k+1) - \mathbf{z}_{t_i}(k)||_2, \tag{2.2}$$

where we note $\mathbf{z}_{t_i}(k)$ the column vector containing the Cartesian coordinates of the detection $z_{t_i}(k)$. If the property 2.2 holds, then extending iteratively each track with the closest detection in the following frames provides the optimal tracking solution. However, these ideal conditions are never met in biological applications and two main deviations from the ideal case are generally observed :

1. The detection $z_{t_i}(k+1)$ is the nearest one for another track than the ith one, and/or there exists a detection $z_{t_j}(k+1)$ that is closer to $z_{t_i}(k)$ than $z_{t_i}(k+1)$. This case is generally due to large particle displacements resulting from a fast target motion compared to the acquisition rate.
2. The detections set is corrupted by the presence of some false detections and some measurements are missing due to the poor image quality. Moreover, when processing biological images, some targets may appear by entering the field of view and the focal plane, while other particles disappear when they leave the observation volume or bleach to a low level of intensity. The number of targets is therefore unknown and time varying.

The described nearest neighbor (NN) principle is not suited to these real conditions since it assumes that the number of targets is known and fixed, and that the detection procedure performs ideally. There exists mainly two simple manners to adapt the NN principle to the problem of tracking a varying number of targets :

- **Local NN (lNN)** : At each frame we build all the possible associations between the set of active tracks and the current set of detections. We select first the association with the smallest distance and the corresponding track and detection are discarded from the association problem. The procedure is then repeated until no more associations are feasible. Remaining detections are used to create new tracks, while non associated tracks are ended.
- **Global NN (glNN)** : We look for the set of associations between tracks at frame k and detections at time $k+1$ with the smallest sum of distances : $\sum_{i=1..n} ||\mathbf{z}_{t_i}(k+1) - \mathbf{z}_{t_i}(k)||_2$. In practice, it is necessary to enumerate all the possible combinations of associations, which may turn out to be time consuming. Track creation and termination are handled with the same *ad hoc* rules than with the lNN.

In order to reduce the computation time, only associations within a distance below d_{max} are allowed. This rule is generally called *association gating*. In practice d_{max} is user-defined and should correspond to the maximum distance traveled by a particle between two subsequent frames. The presence of spurious detections is accounted for by discarding tracks composed of a single detection.

Along the thesis we have experienced that lNN and glNN perform well in high acquisition rate conditions. Indeed, in this case the ambiguity of the tracking is low since detections in subsequent frames are almost overlapping. However, as soon as the acquisition procedure slows down, the tracking performance dramatically drops down for both methods because track switches between targets are done. It indeed often occurs that the closest detection to the previous target location comes from another target when the movements are fast as compared to the acquisition rate. Moreover, NN approaches are unable to deal with poor imaging conditions : the methods cannot discriminate between false and true detections, hence they produce many false tracks. In this challenging case, the inability to deal with missing detections results in an increase number of incomplete tracks.

Summary

The NN approach is the most simple and intuitive method for tracking particles over time. However, the only benefit of such an approach for particle tracking is its easy implementation (which is the reason of its presence in most commercial image processing software). Indeed, essential characteristics of the particle tracking problem are ignored : false and missing detections are not accounting for, while particle dynamics is not modeled. In practice this solution is thus rarely feasible : the image SNR has to be high to avoid missed and false detections, while spots must be nearly overlapping in subsequent frames in order to alleviate the ambiguities of association between detections and tracks.

2.1.2 The point correspondence problem for biological particle tracking

Tracking multiple points is an issue that is often met in vision and surveillance applications, hence, a wealth of algorithms have been dedicated to this problem [150, 42, 87, 137]. Most of these methods belong to a class of algorithm which aims at solving the point correspondence problem : tracking is done by finding correspondences between subsequent sets of detections. Some of these techniques provide solutions to the issues of tracking biological targets in fluorescence images. As an example, glNN and lNN fail to track objects with fast and directed motion, but Sethi and Jain [150] have proposed to modify the distance criterion in order to track objects that move in such a fast and regular way. The association cost function for each track is computed at time k by taking into account two previous positions instead of only one :

$$\begin{aligned} c_i(k) =& \gamma_1 \left[1 - \frac{(\mathbf{z}_{t_i}(k-1) - \mathbf{z}_{t_i}(k-2))(\mathbf{z}_{t_i}(k) - \mathbf{z}_{t_i}(k-1))}{||\mathbf{z}_{t_i}(k-1) - \mathbf{z}_{t_i}(k-2)|| \times ||\mathbf{z}_{t_i}(k-1) - \mathbf{z}_{t_i}(k-2)||} \right] \\ &+ \gamma_2 \left[1 - 2 \frac{\sqrt{||\mathbf{z}_{t_i}(k-1) - \mathbf{z}_{t_i}(k-2)|| \times ||\mathbf{z}_{t_i}(k-1) - \mathbf{z}_{t_i}(k-2)||}}{||\mathbf{z}_{t_i}(k-1) - \mathbf{z}_{t_i}(k-2)|| + ||\mathbf{z}_{t_i}(k-1) - \mathbf{z}_{t_i}(k-2)||} \right]. \end{aligned} \tag{2.3}$$

By doing so smooth tracks are favored : the first term in the cost function c_i favors tracks with constant directions, while the second term favors tracks with regular displacements. γ_1 and γ_2 are tuning parameters between the two regularization terms.

Feature point tracking (FPT)

The development of tracking algorithms for computer vision has inspired some works in the bioimaging field such as the particle tracking algorithm proposed in [147] which is based on the feature point tracking method [42]. We also call feature point tracking (FPT) the algorithm described in [147] in the remainder of the manuscript. In [147] the main adaptation of the original FPT to biological images is the extension of the original point matching algorithm to account for several frames and particle appearance/disappearance events.

In the FPT tracks are represented with association matrices between set of detections. An association matrix G_r^k is build for every pair of detections sets $\{Z(k), Z(k+r)\}_{k=1..l-R}$: $G_r^k(i,j)$ is 1 if $z_i(k)$ is produced by the same physical particle than $z_j(k+r)$, and 0 otherwise. Here $r = 1..R$, with R being a user-defined parameter fixing the number of future frames to account for. Each association matrix is augmented with both a row and a column for virtual particles. Linking a particle to a virtual detection corresponds to an occlusion event of the object, and linking a virtual measurement to a detection reflects the appearance of a new target.

The proposed implementation to find the optimal association is based on the transportation problem [84] which requires the score function of association Φ to be linear in the association variables. The score function to minimize, Φ, is written as a linear combination :

$$\Phi = \sum_{i=0}^{m_k} \sum_{j=0}^{m_{k+r}} \phi_{ij} g_{ij}, \tag{2.4}$$

where m_k and m_{k+r} are the number of detections at time k and at time $k+r$, and the factor ϕ_{ij} is the cost of associating the detection $z_i(k)$ to the detection $z_j(k+r)$. In the original FPT algorithm [42] the smoothness enhancing score [150] is used for the association cost ϕ_{ij}. However, when accounting for several frames, this cost function of association prevents any factorization of the score function Φ. Indeed, each cost ϕ_{ij} at a given time depends on the previous associations. In [147] the following distance is thus used instead :

$$\begin{aligned}\phi_{ij} &= ||\mathbf{z}_i(k) - \mathbf{z}_j(k+r)||_2^2 \\ &\quad + (m_0(z_i(k)) - m_0(z_j(k+r)))^2 + (m_2(z_i(k)) - m_2(z_j(k+r)))^2,\end{aligned} \tag{2.5}$$

which accounts for the Euclidian distance and the difference of intensity moments of order 0 (m_0) and 1 (m_1) between the subsequent detections.

The proposed algorithm is well suited for tracking particles moving slowly compared to the acquisition rate, as shown with the experiments done in [147]. However, like most NN methods, the proposed score function is not adapted to particles moving in a directed way since trajectory smoothness is not evaluated. The introduction of appearance features such as two intensity moments can help solving the association problem in this case. However, the function Φ mixes three terms of different types in a scale dependent way. The scale dependency is however a major issue. Indeed, different intensity normalizations or different settings of the acquisition device would yield different tracking scores because the scale of the intensity values would have changed. Moreover, the confidence in each information used to compute Φ is not addressed in [147]. Indeed, the same weights are given to the intensity fluctuations of a particle with a very stable photo-emission and a particle with an erratic emission (such as QDs). The combination of these issues can lead to incoherent scores and wrong associations. Addressing the issues of information normalization and quality assessment is thus required when designing tracking algorithms for fluorescence images.

Partial tracks linking

In [94] a new tracking procedure has been proposed to track particles in live-cell image sequences. The approach relies on two steps : (1) linking detected particles between consecutive frames ; (2) linking partial track segments generated in the first step to simultaneously close gaps and capture particle merge and split events.

The first linking step involves solving a standard frame-to-frame NN problem : the algorithm is based on a linear formulation similar to the one of the FPT [147] where the association cost is defined as the squared Euclidian distance between detections. However, contrary to the FPT, the tracking algorithm allows only associations between detections in subsequent frames. Indeed, in [94] tracks cannot use virtual measurements to model temporary target disappearance. On the contrary, only partial tracks made of consecutive detections are built, while gaps are expected to be compensated for in the second step of the procedure.

In the second step potential tracks are combined to recover complete tracks. To do so, six links between tracks segments are considered to model track start, end, split, and merge events. First, the end of a track could link to the start of another track, which corresponds to closing a gap due to a target temporary disappearance. Particle merging events are also accounted for by allowing the end of a track to link to a middle point of another partial track. Inversely, a splitting event corresponds to linking a middle point of a track to the start of another one. Starts and ends can also link to nothing, which corresponds to the appearance and the disappearance of a target, respectively. The score of a set of final tracks is computed as the sum of the cost of each of these linking events. Hence, the linking task can be seen as a standard linear assignment problem [31], which makes it fast to solve in a global way [99].

In [94] it is shown that the lifetime of endocytic clathrin coated pits are regulated by dynamin by means of tracking these objects through time. The tracking algorithm was also used to show that the aggregation activity of CD36 receptors depends on their motion type. Despite these two successful applications, the approach may suffer from some limitations in numerous biological cases. Indeed, the second tracking step combines track segments built in the first step, but cannot modify them. Tracking errors thus propagate through steps. The tracking performance therefore critically depends on the first step which is a NN frame-to-frame tracking procedure and thus suffers from similar problems as most NN methods. For instance, false detections are not addressed by the procedure and the temporal information is used only to a very small extent since targets dynamics is not accounted for. Hence, false detections may be assigned instead of true detections, and wrong associations may occur, which results in wrong tracks. Moreover, other issues come from the arbitrary choice of linking costs in the second tracking step. Indeed, events such as target appearance and disappearance, or tracks merging and splitting, should be tailored to the biological application, but these questions are not addressed in [94], as for the tracking of particles moving in a non Brownian way.

2.1.3 Summary

Point correspondence approaches for tracking have been the first methods to be used in biology in an automatic manner [5] due to their simplicity and the existence of numerous previous works in other fields. Such tracking algorithms are nowadays implemented in most image analysis softwares because of their generality. However, we have shown here that point correspondence techniques are useful in a very limited range of applications : the acquisition rate must be high compared to targets movements and the image quality has to be good. Unfortunately, these conditions are rarely satisfied when imaging nano-metric scale particles. Using such methods outside the range of these assumptions would thus result in many false tracks. Most importantly, downstream analyses and interpretations would suffer from a strong bias since only bright and slow particles are correctly tracked. The use of NN methods should therefore be limited to the very few applications which satisfy the strict acquisition rate and image quality requirements.

2.2 Tracking as global optimization problem

Methods that do not use a frame-to-frame detection correspondence have been proposed to overcome the problem of false and missing detections. The key idea is to perform simultaneously particle detection and particle tracking to exploit further the temporal information and the smoothness of trajectories, whereby the poor quality of images can be compensated for. In the following, we browse through the main methods exploiting this approach for processing fluorescence microscopy images.

2.2.1 Single particle trajectory extraction by dynamic programming

In [143] the ill-posedness of the particle segmentation task has been identified as the major weakness of the frame-to-frame approach for tracking : missed and false detections are not simply accounted for by point correspondence algorithms. This limitation is emphasized by the restricted vision in time of frame-to-frame linking procedures. These observations have led the development of a method that takes the exact opposite way : particle detection and tracking are performed simultaneously by analyzing the whole spatio-temporal data set. This approach allows one to extract reliably a particle's trajectory in an extremely noisy 3D image sequence.

A three stages image analysis procedure has been proposed to extract the trajectory of a single particle in a 3D noisy sequence of fluorescence image in [143] :

1. *Sequence alignment.* This module compensates for the movement of the biological sample. The procedure relies on the segmentation of background structures and their translational alignment in the whole sequence.
2. *Spot enhancing.* A pre-filter enhances particle like shapes and reduces the background and the effect of noise : the whitened Laplacian of a Gaussian filter is used (see Chapter 1).
3. *Tracking in time.* The construction of the trajectory of a single particle from the first to the last frame is achieved. The position in each frame is found without explicit detection procedure by directly finding an optimal path in the spatio-temporal volume.

The procedure is designed for the tracking of a single particle which presence is certain over the full length of the noisy image sequence. For one particle the tracking problem is formulated as a global optimization problem with a cost function $\xi(\mathbf{z}^l)$ expressed as :

$$\xi(\mathbf{z}^l) = \sum_{k=1..l} (1-\lambda) f(\mathbf{z}(k), k) - \lambda \frac{||\mathbf{z}(k) - \mathbf{z}(k-1)||_2}{M} \ . \tag{2.6}$$

The functional $f(\mathbf{z}(k), k)$ gives the response of the image to the enhancing filter, at position $\mathbf{z}(k)$ and time k. Its role in Equation 2.6 is to favor positions where the response to the filter is high. The score function $\xi(\mathbf{z}^l)$ mixes this image data with a distance constraint expressed as the Euclidian distance between subsequent positions. Thus, the best path is short and goes through bright pixels. In Equation 2.6, M is a normalization constant and λ is a user defined parameter balancing the image information and the distance traveled by the particle.

The cost function $\xi(\mathbf{z}^l)$ can be computed iteratively from time 1 to l. Moreover, each iteration of the computation scheme involves only the current position $\mathbf{z}(k)$ and previous position $\mathbf{z}(k-1)$ under investigation. Hence, enumerating all possible paths is not necessary, instead the problem is solved by dynamic programming (DP) : the algorithm applies an iterative update formula to compute the maximal cost for reaching the current state. Once the final frame is processed, the state $\mathbf{z}^\star(l)$ with the best score is selected and the whole trajectory is built by back-tracking.

The complete procedure for particle tracking proposed in [143] was validated on synthetic images in low SNR conditions and shown to outperform commercial software solutions based on the frame-to-frame linking approach [90, 174]. The computational time is shown to be short in 2D and manageable in 3D thanks to the DP scheme and the bounding of the displacement between subsequent frames.

The algorithm proposed in [143] can be used for tracking multiple particles only if they are widely apart : in this case a region is delimited around each particle and the algorithm is applied independently for each region. Moreover, dealing with target appearance/disappearance events is not feasible. These constraints considerably restrict the field of application of the method. The approach is however useful in a range of applications in which the target is unique. As an example, the algorithm described in [143] has been used for the study of the chromosome dynamics in budding yeast. The telomere of the right arm of chromosome 6 was specifically labeled by GFP and supposed brighter than other structures that were not labeled. The telomere is always present in the bacteria, which justifies the use of a method that extracts a single trajectory over the whole sequence.

The conditions of use of the method proposed in [143] are not met in the vast majority of biological applications : multiple spots have to be tracked, some entering and leaving the field of view. In such cases the proposed method is not appropriate since handling targets ambiguity would result in a combinatorial explosion of the tracking possibilities. It would thus dramatically increase the cost of the DP technique. The approach also suffers from a common limitation with NN methods : it assumes that particles exhibit a small displacement between subsequent frames, while particles can move fast compared to the acquisition rate. The Euclidian distance criterion is indeed not suited to particles moving in a linear fashion.

2.2.2 Perceptual grouping in the whole spatio-temporal space

In [25] the dynamics of QD-tagged membranar receptors is investigated with particle tracking tools in sequences of 2D images. The QDs used for this study are subject to blinking, so a given QD may be detected during extended periods of time, and, during some time intervals its intensity is too low to be detected. The proposed approach is therefore to use a NN like technique when a particle is easy to detect, and then use a more global approach to compensate for the temporary disappearance events. The two steps are the following :

1. *Building partial tracks with a frame-to-frame approach.* Particles positions are first detected in the whole sequence with a template matching approach (see Chapter 1). The detections are then grouped in partial trajectories with a gated NN frame-to-frame approach.
2. *Trajectories completion by minimal paths.* Partial tracks are assembled to form complete trajectories : tracks are selected and missing positions between the segments are recovered using a minimal path reconstruction in the spatio-temporal volume.

These two steps require the definition of a specific metric to measure distances between QD positions. In [25] the chosen metric is image dependent : it favors short paths that go through high intensity pixels. The corresponding image potential $P(x, y, k)$ is the following :

$$P(x, y, k) = \frac{1}{I_k(x, y) + 1}, \tag{2.7}$$

at position $\{x, y\}$ and time k. The distance along a path linking two points is defined as the integral of the potential $P(x, y, t)$ along it.
We explain now briefly each of the two tracking steps.

Building partial tracks with a frame-to-frame approach

After spot detection, the correspondence problem is solved with a NN tracking approach which is adapted to tracking QD-tagged membranar receptors. At the plasma membrane, early models postulates that proteins move by free diffusion in a lipid bi-layer, hence, the most general assumption is that protein dynamics is a planar Brownian motion [146]. For this motion the maximal distance covered by the particle with probability Ψ during the time-lapse τ can be shown to be :

$$r_\Psi(\tau) = \sqrt{4 \log(1 - \Psi) D\tau}, \tag{2.8}$$

where D is the diffusion coefficient of the motion. The authors propose to use this maximal distance to define only few possible associations between detections at different times, which reduces the linking task complexity. A set of partial tracks is built in a gated NN manner, with respect to the image dependent metric (Eq. 2.7) between detections and the maximal distance criterion $r_\Psi(\tau)$. The tracks are ended as soon as no more detections are found.

Trajectories completion by minimal paths.

Some partial tracks, resulting from the initial linking step, may correspond to a common target which has was not detected in a number of frames. The goal of the second step of the tracking procedure is therefore to group partial tracks with respect to the target to which they correspond, and find missing positions between them. In [25], the 2D time-lapse sequence is considered as a single 3D spatio-temporal volume in which tracks can be viewed as 3D smooth curves. The goal of the tracking procedure is to find the minimal paths between partial tracks which are defined as geodesics in the Riemannian image metric P. The reconstruction of minimal paths is based on the active contour (or *snake*) framework [102]. Given two points, the goal of the active contour is to find the curve $\mathcal{C}$ which minimizes the following energy :

$$E(\mathcal{C}) = \int_{\mathcal{C}} \{w + P(\mathcal{C}(s))\} ds \tag{2.9}$$

where s represents the arc-length parameter and w is a regularization term that avoids the convergence on local minima. The solution of the energy minimization problem is obtained through the computation of the minimal action map $\mathcal{U}$ giving the minimal energy integrated along a path between a starting point p_0 and any given point of the plane. The map $\mathcal{U}$ is computed with a *fast marching method* [151, 44] that proceeds by simultaneously expanding wavefronts from the two points to link until the two fronts meet. The minimal path is found by back-propagation in both directions from the meeting point, which makes the method efficient since the construction of the whole map $\mathcal{U}$ is not required. The back-propagation is achieved by gradient descent along $\mathcal{U}$.

For tracking purposes wavefronts are propagated from every start and end of partial track. The first shock between wavefronts originating from a start and a end indicates the two partial tracks to link by following the minimal path. Theses two tracks are then removed from the problem and the procedure is repeated until no more paths are feasible between remaining starts and ends.

Summary

Extracting a trajectory as a smooth curve in the spatio-temporal volume gives the ability to track particles in very low SNR images, as shown in [143]. However, this approach is not appropriate to track numerous and closely spaced targets since other targets presence is not accounted for when extracting a track. On the other hand two steps frame-to-frame approaches efficiently deal with multiple targets by imposing a strict competition for the detections. The idea exploited by the approach proposed in [25] is thus to take the best of both worlds : use a point correspondence method to track numerous particles when they produce a high level of intensity, and use a minimal path extraction method when the particle intensity is too low to be detectable. The proposed algorithm has been used to process 2D sequences of fluorescence images which were part of experiments on the membrane dynamics of single glycine receptors in live cell cultured neurons. It was shown to successfully recover positions of QDs despite intensity fluctuations and the fluorescence intermittency of the particles. However, simulations in various conditions have shown that this novel approach still suffers from some limitations shared with NN methods. For instance, the tracking performance quickly degrades with an increase of the number of spots because the algorithm cannot discriminate efficiently between closely spaced targets. Moreover, when the acquisition rate decreases, the assumption that trajectories appear as smooth curves in the spatio-temporal volume may become erroneous. The image metric based on the potential P is indeed not suited for modeling very erratic and changing motion types in slow acquisition conditions. Moreover, the computational cost of fast marching methods in 4D spatio-temporal volumes is still an open issue for 3D time-lapse sequence processing.

2.2.3 Kymograph analysis

The principle of the kymograph analysis is to project a multidimensional time-lapse sequence in a lower dimensional space where trajectories visualization and tracks construction are facilitated. Standard kymographs consist in a two dimensional representation of spatio-temporal intensity fluctuations along a particular line $\mathcal{L}$. The line $\mathcal{L}$ is defined in the original volume which can be N-dimensional, but it forms a one dimensional space indexed by the coordinate d along it. Generally the two dimensional kymograph L corresponding to the line $\mathcal{L}$ is defined such that $L(d,t)$ gives the intensity at position d on the line $\mathcal{L}$ at time t. In biology, kymographs have been used to determine global trends of object dynamics along a straight line in 2D [103, 15] and to measure microtubule and actin filament dynamics [180, 81].

More recently, the kymograph analysis has been used for studying the trafficking of the Rab6A protein along microtubules [136] and for estimating the dynamics of microtubules [89]. In [136] authors propose to use kymographs along 3D paths to visualize and quantify Rab6A vesicle motion even if these paths are not *a priori* known, nor specifically labeled. The main assumption founding the technique is that vesicles are transported along microtubules which can be viewed as 3D smooth curves. The procedure therefore consists in two distinct steps : first microtubules are identified in the 3D volume, and then a set of kymographs are built from microtubules detection. Each kymograph is analyzed to extract the tracks of the vesicles.

Kymograph extraction

In [136] authors assume that vesicles move along static microtubules. The first tracking step is thus the localization of these microtubules. If vesicles are numerous and move sufficiently slowly, they are expected to leave a nearly continuous fluorescent trail on the time axis projected data, which corresponds to the microtubule location. The detection of this trail being computationally expensive in the 3D time projected space, a *maximum* projection is applied beforehand along the z-axis, and then along the time-axis. After the projections, the microtubules should appear as 2D smooth curves which are detected by either two methods : (1) semi-automatically with the intervention of the user to determine microtubules extremities and an automatic path optimization to link these points [44], or (2) automatically by the adaptation of a line detection method [172]. For the automatic procedure a number of ambiguous cases have to be resolved, especially when vesicles and microtubules are numerous. For instance when crossroads are detected, lines are split to avoid false connections. Unfortunately, this also prevents the construction of long microtubules.

Once a curve is extracted, the missing vertical coordinate of the microtubule is determined by extracting a 2D curvilinear cross section in the time projected volume along the curve. The resulting line $\mathcal{L}$ in the 3D volume is then used to extract a 2D kymograph L from the original 3D time-lapse sequence.

Despite the simplification brought by projecting 3D+t images in a 2D space the method is still reported to be time consuming in [136] : in practice images are processed by groups of 20 stacks to lower the computational load, and tracks between these groups are linked in a post processing step.

Vesicle tracking

A kymograph L contains the information of intensity fluctuations over time along a given microtubule. Therefore, a fluorescent object moving along $\mathcal{L}$ generates a continuous curve on L if it is slow compared to the acquisition rate, or a dashed line if it is fast. Tracks are therefore extracted from L with a line detection method similar to the one used for microtubule extraction. Tracks defined over less than four consecutive images are considered as originating from the noise and they are therefore removed. Slow tracks (with less than five pixels between the start and end detection points) are considered as originating from the background which is assumed to be nearly static.

Results

The kymograph analysis procedure proposed in [136] was proved efficient for the analysis of Rab6A vesicle trafficking on microtubules in 40 cells : for each cell 3D images where acquired and processed with the fully automatic tracking procedure, which resulted in several hundreds reconstructed tracks for each cell. The efficiency of the proposed method in the Rab6a vesicle tracking application is however counterbalanced by its limitation to few biological applications. The kymograph analysis approach for tracking is indeed based on the assumption that particles move in a directional way along a smooth curve. It is also supposed that targets move with a nearly constant velocity, which allows approximating tracks by lines in the kymographs. These assumptions are however not suited to freely diffusing particles which have a much more erratic motion and do not produce smooth paths in the spatio-temporal volume. Moreover, diffusing particles do not go through common paths, so, one kymograph would be required for each particle. It is worth noting that using this procedure outside the range of its applicability can lead to severe corruption of downstream analyses and erroneous biological interpretations : the method is designed to follow directed particles with nearly constant velocity, so built tracks tend to show these characteristics, while objects that have other motion, such as the Brownian motion, are incorrectly tracked.

2.2.4 Summary

Analyzing time-lapse sequences as a whole constitutes a step further in the improvement of tracking methods. Some important issues, from which early frame-to-frame approaches suffer, have been resolved : in [143] a novel approach for tracking a particle in very low SNR conditions have been presented, and in [25] an algorithm to track intensity fluctuating targets was described. However, the presented algorithms in this part all suffer from a lack of generality. Indeed, the algorithm proposed in [143] is applicable only if multiple targets are widely apart, the acquisition rate have to be slow for the methods proposed in [25] to work properly, while the kymograph analysis approach [136] requires targets to move in a directed way with a constant and low velocity. The use of the presented tracking algorithms therefore imposes strict acquisition constraints, which are often impossible to satisfy in practical biological applications. In contrast we think that analysis methods should adapt to the data, even if they are noisy, and when the acquisition rate is low.

2.3 Conclusion

In this chapter we have browsed through a number of tracking methods proposed for particle tracking and we have shown that point correspondence methods generally suffer from a lack of robustness to poor imaging conditions : missed and false detections are not simply accounted for in a proper way. Hence, only high SNR images can be reliably analyzed. On the other side, processing the time-lapse sequence as a whole can solve this issue of frame-to-frame tracking approaches : poor imaging conditions data sets can be processed by integrating the image and dynamic information from several frames. A number of issues are however yet unresolved such as detecting target appearances and disappearances, and avoiding track coalescence. Some regularity assumptions on particle motion are also generally made, which are fulfilled only with a fast acquisition procedure compared to the particle motion.

The methods presented in this chapter are not appropriate to most biological applications : simple point correspondence methods are unable to account for the specificity of particle tracking in microscopy images, while more complex methods are generally tuned to one specific application and are difficult to generalize to the great variety of biological images. Most importantly is the possible corruption of downstream analyses when using such unadapted algorithms : tracking only very bright spots with a point correspondence method can completely obliterate a biological phenomenon linked to the low intensity particles which cannot be tracked. Moreover, assuming some regularity on track shapes can lead to a severe bias when studying dynamics of particles since tracks fulfilling the regularity assumptions are favored. We therefore argue that the flexibility of tracking methods and the capability to adapt to the data set characteristics are of prime importance for biological applications.

Chapter 3

Probabilistic tracking of biological particles

3.1 Movement statistics for spots linking

Thanks to the development of video microscopy over past decades single particle tracking is now a reference tool in biophysics for the study of membrane proteins mobility [152] and colloidal suspensions [49]. In the biophysics community a considerable work has been devoted to the analysis and modeling of movements of such particles by the mean of tracking experiments [146]. It was shown that in most cases particles diffuse randomly in a 2D plane or in a 3D volume, and several motion subtypes were identified, such as the Brownian motion, the constrained Brownian motion, etc [135]. The random nature of particle movements has led to describe them with statistical models of motion (see the work of Micheal J. Saxton and colleagues [146]) and the idea to exploit these statistical models to track particles in a more reliable way has consequently emerged.

3.1.1 LDL and virus receptor tracking

In [4] a study of low-density lipoprotein (LDL) and influenza virus receptor mobility by means of digital time-lapse microscopy is presented. The motility investigation is divided in five consecutive steps : labeling targets with fluorescent probes, time-lapse acquisition with a fluorescence microscope coupled to a CCD camera, spot position and intensity determination, spot motion tracking, and trajectories analysis. The spot detection procedure is similar to the Gaussian mixture fitting technique described in Chapter 1, relying on the approximation of the LDL and virus particles shapes by the microscope's PSF. The detections linking step is similar to the standard lNN approach introduced in Chapter 2, however, authors propose in [4] to use the probability of moving from one detection to the following one as a cost function (instead of using the Euclidian distance between spots). The probability P of moving from the detection $z(k-1)$ to the location $z(k)$ is expressed as a conditional probability $P(z(k)|z(k-1))$, and computed as the combination of two probabilities :

$$P(z(k)|z(k-1)) \propto \exp\{-(\Delta R/\Delta R_d)^2\} \times \exp\{-(\Delta I)^2/I_n\}, \tag{3.1}$$

where $\Delta R = ||z(k-1) - z(k)||_2$ is the distance of the position change, R_d is a measure of the expected displacement length, $\Delta I = |I_{k-1}(z(k-1)) - I_k(z(k))|$ is the change of the peak intensity, and I_n is the mean peak intensity of the spot. The spatial displacement probability is derived from a normal pdf, which is based on the random diffusion assumption for macro scale time steps [134]. In Equation 3.1 the spatial displacement information is mixed with statistics of intensity change between consecutive spots. The use of the probabilistic framework for the tracking score function allows here mixing information of different types. Moreover, the reliability of each feature is quantified in the score function by the use of a proper pdf : a very robust and accurate feature have a very peaked pdf, while a feature with low reliability has a diffuse pdf which lowers its role in the target discrimination. The information quality assessment is therefore addressed here, in contrast with standard tracking algorithms using non kinetic features to discriminate particles [147, 150, 94].

3.1.2 Tracking colloidal particles in suspension

The probability of particle movements has been used in [49] for tracking individual colloidal particles in time lapse image sequences acquired with a CCD camera mounted on a fluorescence microscope. Five stages for colloidal particle tracking in an image sequence are used : image restoration, particles localization, location refinement, noise discrimination and z position estimation, linking locations into trajectories, and error estimation. Particles location in the whole sequence are obtained by intensity histogram segmentation and gray-scale dilatation [93]. The spot coordinates are further refined by the computation of the brightness-weighted centroid at the corresponding location.

The spots linking procedure explicitly relies on the dynamics of particles which are assumed to be non-interacting and Brownian. For an ensemble of n particles the displacement probability from a set of positions $Z(k-1)$ to the positions $Z(k)$ is written as the combination of the normal probability values of the displacements :

$$P(Z(k)|Z(k-1)) = \Big(\frac{1}{4\pi D\Delta t}\Big)^n \exp\left(- \sum_{i=1..n} \frac{||\mathbf{z}_i(k) - \mathbf{z}_i(k-1)||_2^2}{4D\Delta t} \right) \tag{3.2}$$

where D is the common diffusion coefficient. It is easy to show in this case that the most likely assignment between $Z(k-1)$ and $Z(k-1)$ minimizes the sum of the squared distances :

$$\begin{aligned} \{\mathbf{z}_i(k)\}^{\star}_{i=1..n} &= \arg \max_{\{\mathbf{z}_i(k)\}_{i=1..n}} P(Z(k)|Z(k-1)) \\ &= \arg \min_{\{\mathbf{z}_i(k)\}_{i=1..n}} \sum_{i=1..n} ||\mathbf{z}_i(k) - \mathbf{z}_i(k-1)||_2^2 \end{aligned} \tag{3.3}$$

This statement shows that results obtained with NN methods using the Euclidian distance do not correspond to the maximum likelihood principle with an underlying Brownian assumption. To do so one should use the squared Euclidian distance instead of the Euclidian distance. Using the square Euclidian distance between subsequent detections penalizes to a greater extent unlikely movements that using the Euclidian distance.

3.1.3 Quantum dots (QDs) tracking

Multi Target Tracking (MTT) is a frame-to-frame algorithm dedicated to quantum dots (QDs) tracking in 2D+t images which was recently proposed in [148]. The approach is very similar to the ones proposed in [4] and [49], however, some important modifications have been made to the computation of the probability of moving from one detection to another in order to account for the blinking property of QDs. The probability for a target to move from one 2D detected position $z(k)$ to the subsequent one $z(k+1)$ is computed as the following combination of three probabilities :

$$p(z(k+1)|z(k)) = p_{diff}(r) \cdot p_{int}(I) \cdot p_{off}(k), \tag{3.4}$$

which accounts for the probability of the movement of the particle ($p_{diff}(r)$), the probability of its intensity ($p_{int}(I)$), and its possible extinction due to blinking at time k ($p_{off}(k)$).

Movement probability

The Euclidian distance covered by the particle between frames k and $k+1$ is noted r, and its probability is computed assuming that particles diffuse in a Brownian way. Two types of diffusion are considered : the local one with a diffusion coefficient D_{local}, and the maximum diffusion with a diffusion coefficient : $D_{max} > D_{local}$, which accounts for sudden increase towards free diffusion. The resulting probability is computed as follows :

$$\begin{aligned} p_{diff}(r) &= \gamma_{diff} p_{local}(r) + (1-\gamma_{diff}) p_{max}(r) && (3.5)\\ &= \gamma_{diff}/(\sqrt{2\pi} r_{local}) \exp(-r^2/2r_{local}^2) \\ &+ (1-\gamma_{diff})/(\sqrt{2\pi} r_{max}) \exp(-r^2/2r_{max}^2), && (3.6) \end{aligned}$$

by weighting the two possible modes of diffusion with the help of the parameter $\gamma_{diff} \in [0,1]$ (set to 0.9 in [148]). The parameters r_{local} and r_{max} in Equation 3.6 correspond to the standard deviations of the pdfs of the Gaussian laws computed for the local and maximum diffusion movements, respectively. The diffusion coefficient D_{local} is continuously updated over time by estimating it within the temporal window $w_k = [k-4, \ldots, k]$, while D_{max} is fixed beforehand. The diffusion coefficient update allows the method to estimate a different pdf of movement for each particle, depending on the length of its recent movements.

Intensity probability

The probability law for the intensity I of a particle is computed as the combination of two pdfs at a given frame k :

$$\begin{aligned} p_{int}(I) =& \gamma_{int} p_{on}(I) + (1-\gamma_{int}) p_{blink}(I) && (3.7)\\ =& \gamma_{int}/(\sqrt{2\pi}\sigma_I) \exp(-(I - <I_{on}>)^2/2\sigma_i^2) \\ &+ U_{[0,<I_{on}>]}(I). \end{aligned}$$

In Equation 3.7 the first term corresponds to the probability law of the intensity of the particle when it does not blink. The particle intensity in this 'on' mode is assumed to fluctuate in a Gaussian way around the mean intensity $< I_{on} >$. In practice $< I_{on} >$ is estimated from the 5 previous observed intensities of the particle in the 'on' state. The second term in Equation 3.7 corresponds to the pdf of the intensity when the particle blinks. In this case the particle intensity is assumed uniformly distributed between 0 and $< I_{on} >$ since an accurate pdf cannot be defined in this case. The functional $U_{[0,<I_{on}>]}(I)$ thus returns $(1-\gamma_{int})/ < I_{on} >$ if $I \in [0, < I_{on} >]$, and 0 otherwise. According to the authors evaluation, both intensity states have comparable probabilities, thus the mixing parameter γ_{blink} is set to 0.5.

Probability law for full blink and disappearance

The term $p_{off}(k)$ corresponds, for a particle that has disappeared in previous frames, to the probability to re-appear in the current image. Three cases are therefore distinguished :

- $p_{off}(k) = 1$ if the particle was associated to a real detection in the previous frame. In this case the particle did not disappeared and was not fully blinked at time $k-1$.
- $p_{off}(k) = \exp(-k - k_{off})/\tau_{off}$ if the particle is fully blinked since the frame k_{off}. The probability of re-appearing has an exponential decay over time with respect to the average blinking duration τ_{off}.
- $p_{off}(k) = 0$ if $k - k_{on} > 3\tau_{off}$. A bound on the duration of the blinked state is applied here : a particle that has not been in the 'on' mode for more than $3\tau_{off}$ frames cannot appear again. In such a case, the corresponding track is ended by setting the association probability $p(z(k+1)|z^k)$ to zero, whatever the detection $z(k+1)$.

The three different probabilities allow modeling visible particles, blinked QDs, and targets that have exited the surveillance volume or have physically disappeared.

Association problem solving and tracks management

In [148] the association of tracks with detections is dealt with in a very same way as in the lNN algorithm : all the associations satisfying a distance bound are built, the links between tracks and detections are then iteratively selected in a decreasing order the association probability $p(z(k+1)|z(k))$. When a track does not find any detection it is associated to a virtual measurement which models a blinking event. Each non associated detection is used to initialize a new track since it is assumed that the detection set does not contain any spurious measurement. The parameters characterizing the movement and the stochastic evolution of the intensity are updated with respect to the selected association.

Summary

In [148] the introduction of an advanced model of intensity fluctuations have permitted tracking numerous blinking QDs in 2D+t images. These tracks have served to the estimation of a spatio-temporal map of membranes in the cell. However, the downsides of the algorithm are similar as for the lNN : the simplistic view of the association problem may result in poor tracking results in noisy and dense environments. Indeed, false detections are not accounted for and events such as simultaneous targets appearance and blinking are neglected. Moreover, the inaccuracies corrupting the measurement process are not addressed, and other motion types than the diffusive one are not accounted for. Hence, the proposed probability computations may be incorrect in frequent biological cases.

3.2 State space filtering

Numerous issues, such as defining proper pdfs and tuning parameters, have to be dealt with when using statistical concepts for particle tracking. As an example, in [4] authors have noticed that the pdf describing the intensity change in the spot transition probability (Eq. 3.1) is unable to model a number of important cases. In particular some changes of focus due to the particle diffusion along the vertical axis may occur, hence leading to abrupt peak intensity changes. The pdf of intensity change is unable to model such fast fluctuations, which makes the true association between the spots very unlikely. Targets motion models have also to be chosen with care : the motion model must be as realistic as possible since it takes part in the computation of the likelihood of association between spots. Dynamics of diffusing particles has been shown to be far more complex than the simple Brownian motion [146, 135], while actively transported particles exhibit very different movements. Defining a complex motion model accounting for the variety of particle motion is therefore required.

We describe next a statistical framework allowing to incorporate in a proper way the wealth of information involved in the particle tracking task.

3.2.1 Bayesian framework for tracking

Bayesian statistical concepts [13] form a flexible framework allowing one to incorporate complex particle motion models, to include other random values such as spot intensity, and to model measurements inaccuracy. Hence, deriving spot linking score functions in a consistent way with known biophysical properties is made possible. Because of these benefits the Bayesian framework is established as a reference way to deal with tracking problems in other fields such as aircraft tracking in radar modalities [11] and in video surveillance (see [48] for a review of statistical methods for solving the motion correspondence problem).

In most tracking applications the computation of various statistics involved in the Bayesian framework is achieved by the use of concepts taken from the broader field called the *estimation theory*[175]. Estimation consists in inferring a value of interest from some observations which can be inaccurate, partial, uncertain and indirect. In tracking applications the value of interest is generally defined as the true position of targets and the observations refer to measurements produced by the used sensor. Bayesian tracking methods assume that the target motion and the observations can be represented by some known mathematical models, which allows computing accurately the pdf of the targets presence in images. The estimated pdfs allow one to build tracks according to various criteria such as the Maximum Likelihood (ML), the Minimum Mean Squared Error (MMSE), and the Maximum-A-Posteriori (MAP) criteria.

The Bayesian framework has been transposed recently to the biological particle tracking problem [64, 154]. In the following, we browse through the main estimation concepts involved in these techniques.

3.2.2 Single track model

The models of target are generally defined in the state *state space* formalism : in discrete time, each target is described at time k by a vector $\mathbf{x}(k)$ of n state variables. The state of a target is assumed to depend on the previous state following the so-called *state evolution equation*, which is written as :

$$\mathbf{x}(k+1) = f_k(\mathbf{x}(k)) + \mathbf{w}(k), \tag{3.8}$$

when an additive noise is assumed. In Equation 3.8, the vector $\mathbf{w}(k)$ is a random process noise at the discrete time k. The vector valued function f_k describes the deterministic part in the evolution from state $\mathbf{x}(k)$ to state $\mathbf{x}(k+1)$, and is possibly time varying.

We note thereafter $\mathbf{x}^k = \{\mathbf{x}(t)\}_{t=1..k}$ and $\mathbf{z}^k = \{\mathbf{z}(t)\}_{t=1..k}$ the sets of states and measurements originating from the target, respectively, from time 1 to time k. We assume that $\mathbf{z}(k)$, the measurement at time k, does not depend on the previous states and measurements, conditional to the current state $\mathbf{x}(k)$:

$$p(\mathbf{z}(k)|\mathbf{z}^{k-1}, \mathbf{x}^k) = p(\mathbf{z}(k)|\mathbf{x}(k)), \tag{3.9}$$

and we link the current measurement to the current state with the *measurement equation* :

$$\mathbf{z}(k) = h_k(\mathbf{x}(k)) + \mathbf{v}(k) \tag{3.10}$$

where h_k and $\mathbf{v}(k)$ are the measurement function and the measurement noise, respectively. The measurement function h_k establishes a deterministic link between the measured features and the state variables, while the measurement noise models the randomness in this relationship.

The estimation of the state vectors from the associated measurements and the computation of the subsequent track probability are fairly common signal processing tasks for which standard estimation techniques exist. The Kalman filter [101] is the preferred technique when both the measurement and the evolution are linear with white and Gaussian noise (WGN). Otherwise, numerous alternative estimation schemes exist. We consider next these two cases separately.

3.2.3 Kalman filtering

When the functions f_k and h_k are linear the evolution and measurement equations can be re-written with the following linear algebraic representation :

$$\mathbf{x}(k+1) = \mathbf{F}_k\mathbf{x}(k) + \mathbf{w}(k) \tag{3.11}$$

$$\mathbf{z}(k) = \mathbf{H}_k\mathbf{x}(k) + \mathbf{v}(k) \tag{3.12}$$

where $\mathbf{F}_k$ and $\mathbf{H}_k$ are the state transition matrix and the measurement process matrix, respectively. The process noise and the measurement noise are assumed white and Gaussian with zero mean and known auto-covariance matrices $\mathbf{Q}(k)$ and $\mathbf{R}(k)$:

$$p(\mathbf{w}(k)) = \mathcal{N}(\mathbf{w}(k); 0, \mathbf{Q}(k)) \tag{3.13}$$

$$p(\mathbf{v}(k)) = \mathcal{N}(\mathbf{v}(k); 0, \mathbf{R}(k)). \tag{3.14}$$

The linearity and WGN assumptions yield a Gaussian distribution of the conditional probability of the observations and states :

$$p(\mathbf{x}(k+1)|\mathbf{x}(k)) = \mathcal{N}(\mathbf{x}(k+1); \mathbf{F}_k\mathbf{x}(k), \mathbf{Q}(k)) \tag{3.15}$$

$$p(\mathbf{z}(k)|\mathbf{x}(k)) = \mathcal{N}(\mathbf{z}(k); \mathbf{H}_k\mathbf{x}(\mathbf{k}), \mathbf{R}(k)). \tag{3.16}$$

The conditional pdfs of the state and the measurement are centered on the predicted state $\mathbf{F}_k\mathbf{x}(k)$ and on the predicted measurement $\mathbf{H}_k\mathbf{x}(k)$, respectively.

From Equation 3.15 and Equation 3.16, the joint state posterior density $p(\mathbf{x}^k|\mathbf{z}^k)$ is also Gaussian and the MAP state estimation problem admits a solution $\mathbf{x}^{\star,k} = \arg\max p(\mathbf{x}^k|\mathbf{z}^k)$ which is also the MMSE solution [13]. In this case, the Kalman filter (KF) [101] is able to retrieve exactly this optimal state $\mathbf{x}^{\star,k}$ in an iterative way by a simple combination of linear filtering and update steps.

At each frame the KF allows estimating the conditional pdf of the state $\mathbf{x}(k)$ as :

$$p(\mathbf{x}(k)|\mathbf{z}^{k-1}) = \mathcal{N}(\mathbf{x}(k); \mathbf{x}(k|k-1), \mathbf{P}_{k|k-1}) \tag{3.17}$$

$$p(\mathbf{x}(k)|\mathbf{z}^{k}) = \mathcal{N}(\mathbf{x}(k); \mathbf{x}(k|k), \mathbf{P}_{k|k}), \tag{3.18}$$

where $\mathbf{x}(k|k-1)$ is the predicted state at frame k when measurements up to time $k-1$ are known, and $\mathbf{x}(k|k)$ is the predicted state with known measurements up to time k. The matrices $\mathbf{P}_{k|k-1}$ and $\mathbf{P}_{k|k}$ are the covariance matrices of the pdfs for each case respectively. The prediction process is achieved by applying the deterministic part of the evolution equation to the previously estimated state $\mathbf{x}(k-1|k-1)$:

$$\mathbf{x}(k|k-1) = \mathbf{F}_k\mathbf{x}(k-1|k-1). \tag{3.19}$$

From the state measurement and the measurement equation it is then possible to compute the expected measurement $\mathbf{z}(k|k-1)$ at time k as :

$$\mathbf{z}(k|k-1) = \mathbf{H}_k\mathbf{x}(k|k-1). \tag{3.20}$$

Once the measurement $\mathbf{z}(k)$ is available the state prediction $\mathbf{x}(k|k-1)$ is corrected to form the estimated state $\mathbf{x}(k|k)$. As shown by the following equation :

$$\mathbf{x}(k|k-1) = \mathbf{x}(k|k) + \mathbf{K}_k(\mathbf{z}(k) - \mathbf{z}(k|k-1)), \tag{3.21}$$

where $\mathbf{K}_k$ is the gain matrix, the estimated state is obtained by adding a term to the predicted state, which depends on the difference between the new observation $\mathbf{z}(k)$ and the predicted measurement $\mathbf{z}(k|k-1)$. This difference is a random value coined the *innovation* of the process. The probabilistic distribution of the innovation is known beforehand and does not depend on the measurements, hence its is usually used to quantify the goodness of fit of the model : accurate evolution and measurement models yield small and probable innovation values.

KF is a powerful estimation technique for state estimation that is remarkably inexpensive for low-dimensional spaces, indeed, it involves only few low dimensional matrix products and inversions. Unfortunately, the simplicity comes with the price of strong requirements : the evolution and measurement process have to be linear with an additive WGN.

3.2.4 Deterministic estimation of non linear and/or non Gaussian models

The KF technique is feasible only for a small class of linear with WGN problems, however, tracking applications are various and KF assumptions are generally not satisfied. For instance, the state transition equation may depart from linearity, hence, nonlinear filtering techniques have to be used in place of the KF in these cases [162, 14, 140]. A main advantage of KF is to give an analytical solution to the state estimation problem, while non linear/non Gaussian systems do not have an analytical solution in general. Hence nonlinear estimation techniques only retrieve an approximate solution of $\mathbf{x}^{\star,k}$. We now list few of these techniques and give a short overview for each one.

Extended Kalman Filter (EKF)

The EKF [14] proceeds in a two steps prediction/update procedure similar as to the standard KF despite the nonlinearity of the state evolution equation or the measurement equation. Indeed, nonlinear functions f_k and h_k are approximated by the first term in their Taylor series expansion, from each expansion the approximated matrices $\hat{\mathbf{F}}_k$ and $\hat{\mathbf{H}}_k$ are computed and used in place of $\mathbf{F}_k$ and $\mathbf{H}_k$ for the estimation of the covariance matrices in the KF scheme.

The EKF and its extensions, such as the iterated EKF and the higher-order EKF, are referred to as *analytic* approximations since linear approximation matrices $\hat{\mathbf{F}}_k$ and $\hat{\mathbf{H}}_k$ are computed analytically from the nonlinear functions f_k and h_k. The EKF is a deterministic estimation method thanks to this analytical approximation. However, a limitation comes from the Taylor series expansion which requires f_k and h_k to be differentiable. Moreover, the EKF approximates the posterior $p(\mathbf{x}(k)|\mathbf{z}^k)$ with a normal distribution, which may be a major issue if the non linearity in the state evolution and measurements equation is severe since the non Gaussianity of the posterior is emphasized in this case.

Unscented Kalman Filter (UKF)

Contrary to the EKF, the UKF [100] does not approximate the nonlinear functions f_{k-1} and h_k. Instead, EKF captures the nonlinearity in the approximation of the posterior density $p(\mathbf{x}(k)|\mathbf{z}^k)$. The main idea behind UKF is to approximate the posterior $p(\mathbf{x}(k)|\mathbf{z}^k)$ by a Gaussian density, which is represented by a set of selected points. These points are chosen in a deterministic manner and propagated through time thanks to a nonlinear function. The use of the unscented transform ensures that the set of points captures the mean and covariance up to the second order of nonlinearity. EKF is therefore based on a statistical linearization instead of an analytical regression.

The UKF is a simple and fast way to approximate the state posterior density, however it is not able to model high-order moments of the true non-Gaussian distribution. This point is critical for some posterior densities which are frequent in tracking applications, such as multimodal densities.

Interacting Multiple-Model filter (IMM)

A possible solution to model a multimodal posterior density $p(\mathbf{x}(k)|\mathbf{z}^k)$ is to approximate it by a mixture of Gaussian densities instead of a single Gaussian density (as done in KF, EKF and UKF). In this case, the number of models, as well as their weights, have to be tuned cautiously to get an accurate estimation of the state posterior density. When the number of models is fixed to a given value p, we note r_k the r-th state at time k, with $r_k \in [1, .., p]$. In this case two main classes of methods exist : those considering that only one filter is correct in the bank of models over the whole sequence, and those considering that the target can switch between models in the manner of a Hidden Markov Model (HMM). In the latter case the transition probability between the state r_{k-1} at time $k-1$ to the state r_k is defined as :

$$\pi_{ij} = P\{r_k = j | r_{k-1} = i\}, \tag{3.22}$$

and generally assumed fixed over time. Assuming a state model can switch between p models yields a exponentially growing number of hypotheses of model sequences through time. Hence, applying such a technique would require efficient pruning techniques to restrict the hypotheses to the most probable ones. Instead, the IMM [21] merges the mixture components, which results in a single mixed hypothesis. The IMM assumes that the input of the filter j at time k is an interaction of all p filters in the bank, hence, a mixing step of all the components is performed before the filtering step of each model. The IMM filter proceeds mainly in three steps when a measurement $\mathbf{z}(k)$ is received at time k :

1. Mixing of the state estimate. For each model j the mixed estimate $\hat{\mathbf{x}}^{0j}(k-1|k-1)$ and covariance matrix $\mathbf{P}^{0j}(k-1|k-1)$ are computed as :

$$\hat{\mathbf{x}}^{0j}(k-1|k-1) = \sum_{i=1..p} \mu^{i|j}(k-1|k-1)\hat{\mathbf{x}}^{i}(k-1|k-1), \tag{3.23}$$

$$\begin{aligned}\mathbf{P}^{0j}(k|k) = \sum_{i=1..p} \mu^{i|j}(k-1|k-1)\Big[&\mathbf{P}^{i}(k-1|k-1) \\ &+ (\hat{\mathbf{x}}^{i}(k-1|k-1) - \hat{\mathbf{x}}^{0j}(k-1|k-1)) \\ &\cdot (\hat{\mathbf{x}}^{i}(k-1|k-1) - \hat{\mathbf{x}}^{0j}(k-1|k-1))^T\Big],\end{aligned} \tag{3.24}$$

with $\hat{\mathbf{x}}^i(k-1|k-1)$ the estimated state for the mode i at time $k-1$, and $\mathbf{P}^i(k-1|k-1)$ its covariance matrix. The conditional probability $\mu^{i|j}(k-1|k-1)$ is defined as :

$$\mu^{i|j}(k-1|k-1) \triangleq P\{r_{k-1} = i | r_k = j, \mathbf{z}^{k-1}\}, \tag{3.25}$$

and is computed by applying the Bayes rule (see proof in [21]) as :

$$\mu^{i|j}(k-1|k-1) = \frac{\pi_{ij}\mu^i(k-1|k-1)}{\sum_{q=1..p}\pi_{qj}\mu^j(k-1|k-1)} \tag{3.26}$$

where $\mu^j(k-1|k-1) \triangleq Pr\{r_{k-1} = j|\mathbf{z}^{k-1}\}$ is the probability of the model j at time $k-1$ conditional to the measurements $\mathbf{z}^{k-1}$.

2. Model-conditioned updates. Each model j is updated with $\mathbf{z}(k)$ by applying a KF using the matrices $\hat{\mathbf{x}}^{0j}(k-1|k-1)$ and $\mathbf{P}^{0j}(k-1|k-1)$. By doing so, $\hat{\mathbf{x}}^j(k|k)$ and $\mathbf{P}^j(k|k)$ estimates are computed.
3. Model probability updates. The model j probability is computed at time k as :

$$\mu^j(k|k) = \frac{p(\mathbf{z}(k)|\mathbf{z}^{k-1}, r_k = j)\sum_{i=1..p}\pi_{ij}\mu^i(k-1|k-1)}{\sum_{q=1..p}[p(\mathbf{z}(k)|\mathbf{z}^{k-1}, r_k = q)\sum_{i=1..p}\pi_{iq}\mu^i(k-1|k-1)]}, \tag{3.27}$$

where the measurement probability $p(\mathbf{z}(k)|\mathbf{z}^{k-1}, r_k = j)$ is computed by the KF of the model j.

At each time point k, the posterior state density is approximated by the weighted mixture of the p Gaussian pdfs given by the KFs of the models :

$$p(\mathbf{x}(k)|\mathbf{z}^k) \approx p_{IMM}(\mathbf{x}(k)|\mathbf{z}^k) = \sum_{i=1..p} w_i^k \mathcal{N}\Big(\mathbf{x}(k); \hat{\mathbf{x}}^i(k|k), \mathbf{P}^i(k|k)\Big). \tag{3.28}$$

The computational cost of the IMM is almost equal to the sum of the costs of the p filters since it fuses the mixture components. This cost is therefore very low when using a mixture of low-dimensional KFs. The capability of the IMM to model multimodal posterior densities and its efficiency have made it very popular for tracking applications : it was used in various fields such as maneuvring target tracking in radar surveillance [12, 181] and in particle tracking applications [68].

3.2.5 Density estimation with Sequential Monte Carlo estimation methods

Sequential Monte Carlo (SMC) methods perform a suboptimal numerical estimation which is generally used for nonlinear and non Gaussian state space models admitting no analytical solution. The estimation is based on the evolving distribution of a set of random point mass of density probabilities. Originally considered as being computationally too expensive, a number of methodological advances and the increased computational power have made SMC estimation methods very popular in a wide range of applications such as in financial econometrics, computer vision, robotics and target tracking (see [54] and references therein).

Sequential Importance Sampling

The Sequential Importance Sampling (SIS) is a general SMC integration that is the base of most recent SMC algorithms used for tracking targets. In the context of filtering, it was also named bootstrap filtering [77, 76], the condensation algorithm [91] and often particle filtering

[7, 140]. The principle of SIS is to approximate the state posterior density $p(\mathbf{x}^k|\mathbf{z}^k)$ by discrete weighted mixture, as follows :

$$p(\mathbf{x}^k|\mathbf{z}^k) \approx p_{SIS}(\mathbf{x}^k|\mathbf{z}^k) = \sum_{i=1..n} w^{k,(i)} \delta(\mathbf{x}^k - \mathbf{x}^{k,(i)}), \tag{3.29}$$

where δ is the Dirac delta mass centered in 0. The state vectors $\{\mathbf{x}^{k,(i)}\}_{i=1..n}$ are called *particles*, and are drawn randomly. In the case of *importance sampling*, particles are drawn from the *importance density* $q(\mathbf{x}^k|\mathbf{z}^k)$ with proper normalized weights $w^{k,(i)}$.

Ideally we would like to sample particles directly from $p(\mathbf{x}^k|\mathbf{z}^k)$ and set the weights equal. In this case the posterior can be computed as :

$$p(\mathbf{x}^k|\mathbf{z}^k) = \sum_{i=1..n} \delta(\mathbf{x}_k - \mathbf{x}^{k,(i)})/n. \tag{3.30}$$

Unfortunately, generating samples directly from $p(\mathbf{x}^k|\mathbf{z}^k)$ is in general not straightforward. Sampling in another distribution $q(\mathbf{x}^k|\mathbf{z}^k)$ and adjusting the particle weights is often much easier, and still makes the MC integration proper. The only requirement being that $q(\mathbf{x}^k|\mathbf{z}^k)$ has the same support as $p(\mathbf{x}^k|\mathbf{z}^k)$. When using the importance density function $q(\mathbf{x}^k|\mathbf{z}^k)$, the expectation of any function $f(\mathbf{x}^k)$ following $p(\mathbf{x}^k|\mathbf{z}^k)$ can be re-written :

$$\int f(\mathbf{x}^k) p(\mathbf{x}^k|\mathbf{z}^k) d\mathbf{x}^k = \int f(\mathbf{x}^k) \frac{p(\mathbf{x}^k|\mathbf{z}^k)}{q(\mathbf{x}^k|\mathbf{z}^k)} q(\mathbf{x}^k|\mathbf{z}^k) d\mathbf{x}^k, \tag{3.31}$$

where the integration is performed on the support of $p(\mathbf{x}^k|\mathbf{z}^k)$ and $q(\mathbf{x}^k|\mathbf{z}^k)$. By drawing n samples : $\{\mathbf{x}^{k,(i)}\}_{i=1..n}$, in the importance distribution $q(\mathbf{x}^k|\mathbf{z}^k)$ the integration in Equation 3.31 can be approximated as follows :

$$\int f(\mathbf{x}^k) p(\mathbf{x}^k|\mathbf{z}^k) d\mathbf{x}^k \approx \sum_{i=1..n} f(\mathbf{x}^k) w^{k,(i)}, \tag{3.32}$$

where the normalized weight $w^{k,(i)}$ is computed as :

$$w^{k,(i)} = \frac{\tilde{w}^{k,(i)}}{\sum_{j=1..n} \tilde{w}^{k,(j)}}, \tag{3.33}$$

based on the unnormalized weight $\tilde{w}^{k,(i)}$:

$$\tilde{w}^{k,(i)} = \frac{p(\mathbf{x}^{k,(i)}|\mathbf{z}^k)}{q(\mathbf{x}^{k,(i)}|\mathbf{z}^k)}. \tag{3.34}$$

Sampling into $p(\mathbf{x}^k|\mathbf{z}^k)$ is alleviated by the importance sampling. As shown in Equation 3.34, one has only to be able to compute the probability $p(\mathbf{x}^k|\mathbf{z}^k)$ point wise.

In time series analysis applications an iterative computation of $p(\mathbf{x}^k|\mathbf{z}^k)$ and $q(\mathbf{x}^k|\mathbf{z}^k)$ is desirable to decrease the computational cost when k is large. While the Markov property of state transition equation (Eq. 3.8) allows one to compute $p(\mathbf{x}^k|\mathbf{z}^k)$ by a succession of updates, the choice of the importance density has to make possible the computation of $q(\mathbf{x}^k|\mathbf{z}^k)$ at any frame k from the value of $q(\mathbf{x}^{k-1}|\mathbf{z}^{k-1})$. To do so, the importance density is generally chosen to factorize as :

$$q(\mathbf{x}^k|\mathbf{z}^k) \triangleq q(\mathbf{x}(k)|\mathbf{x}^{k-1}, \mathbf{z}^k) q(\mathbf{x}^{k-1}|\mathbf{z}^{k-1}), \tag{3.35}$$

such as the samples at time k are obtained by augmenting each particle $\mathbf{x}^{k-1,(i)}$ with a new state $\mathbf{x}^{(i)}(k)$ which is drawn from $q(\mathbf{x}^{(i)}(k)|\mathbf{x}^{k-1,(i)}, \mathbf{z}^k)$. In this case, it can be shown that the weights $w^{k,(i)}$ at frame k are obtained by a simple update of the previous weights $w^{k-1,(i)}$:

$$w^{k,(i)} \propto w^{k-1,(i)} \frac{p(\mathbf{z}(k)|\mathbf{x}^{(i)}(k))p(\mathbf{x}^{(i)}(k)|\mathbf{x}^{(i)}(k-1))}{q(\mathbf{x}^{(i)}(k-1)|\mathbf{x}^{(i)}(k-1), \mathbf{z}(k))}. \tag{3.36}$$

In Equation 3.36 the update factor explicitly accounts for the agreement of the new state $\mathbf{x}^{(i)}(k)$ with the new measurement $\mathbf{z}^{(i)}(k)$ by including the probability $p(\mathbf{z}(k)|\mathbf{x}^{(i)}(k))$ in the computation. When choosing the importance density and the importance weights as defined by Equation 3.35 and 3.36, respectively, it can be shown that the SIS posterior density (Eq. 3.29) tends to the true posterior $p(\mathbf{x}^k|\mathbf{z}^k)$ when the number of particles goes to infinity.

SIS algorithms are appealing techniques for state estimation in non linear/non Gaussian systems because of their flexibility and simplicity. However, simple SIS algorithms suffer from some important issues which attracted lots of attention over last decades.

Degeneracy and sample impoverishment problems

It has been shown that with the basic SIS scheme the variance of importance weights strictly increases over time [55], which indicates that after few SIS steps all the particles, except one, generally have negligible weights. This issue comes from the spread of the particles in the space to sample in, the dimensionality of which increases each time a new time point is considered. Hence, when k is large most samples $\mathbf{x}^{k,(i)}$ have nearly zero probability $p(\mathbf{x}^{k,(i)}|\mathbf{z}^k)$. The inconvenience of the degeneracy of the particles set are twofolds : lots of computational power is devoted to particles that are negligible, and the posterior density is approximated by only one particle which reduces the estimation accuracy.

To prevent an excessive spread of particles it was proposed to re-sample particles to decrease the variance of the importance weights when the degeneracy of the particles (measured by the effective sample size [106]) is too high. The re-sampling procedure generally consists in eliminating particles with too low weights and replacing them with duplicates of high weight particles. The loss of diversity decreases the importance weights variance, however it can also results in a *sample impoverishment* problem. Indeed, in some cases, few particles rapidly dominate the set and eliminate alternative possibilities over time. Sample impoverishment is therefore a major issue for the estimation of multimodal densities which requires sufficient particle spread over the space.

Another popular way to reduce degeneracy is to use Monte Carlo Markov Chain (MCMC) step schemes [73], which are often based on the Metropolis-Hastings algorithm [141]. The main idea is to re-sample each particle with a probability α, which is called the *acceptance probability* of a move. The probability α is computed differently for each particle $\mathbf{x}^{k,(i)}$, depending on the re-sampled particle $\mathbf{x}^{k,(i\star)}$ which is candidate for its replacement. The α value is given by :

$$\alpha = \min \left\{ 1, \frac{p(\mathbf{z}(k)|\mathbf{x}^{(i\star)}(k))p(\mathbf{x}^{(i\star)}(k)|\mathbf{x}^{(i)}(k-1))}{p(\mathbf{z}(k)|\mathbf{x}^{(i)}(k))p(\mathbf{x}^{(i)}(k)|\mathbf{x}^{(i)}(k-1))} \right\} \tag{3.37}$$

for the Metropolis-Hastings algorithm. The MCMC guarantees an asymptotic approximate of the posterior density $p(\mathbf{x}^k|\mathbf{z}^k)$ and it has been shown to alleviate the problem of sample impoverishment. Being not significantly more computationally expensive than other re-sampling schemes, its use has spread over the past years, especially in target tracking applications [104].

Importance density choice

The choice of $q(\mathbf{x}^k|\mathbf{x}^{k-1,(i)},\mathbf{z}^k)$ is of prime importance when designing a SMC method since it constitutes the core of the sampling strategy. The only requirement for Equation 3.31 to hold is that importance density has the same support as $p(\mathbf{x}^k|\mathbf{x}^{k-1},\mathbf{z}^k)$. An intuitive choice for the importance density $q(\mathbf{x}^k|\mathbf{x}^{k-1,(i)},\mathbf{z}^k)$ is therefore $p(\mathbf{x}^k|\mathbf{x}^{k-1},\mathbf{z}^k)$ itself. It has been shown in [55] that $p(\mathbf{x}^k|\mathbf{x}^{k-1},\mathbf{z}^k)$ is the optimal choice since it minimizes the variance of the importance weights and therefore alleviates the degeneracy problem as much as possible. When using $p(\mathbf{x}^k|\mathbf{x}^{k-1},\mathbf{z}^k)$ as the importance density function, the importance weights $w^{k,(i)}$ are computed as :

$$w^{k,(i)} \propto w^{k-1,(i)} p(\mathbf{z}(k)|\mathbf{x}^{(i)}(k-1)), \tag{3.38}$$

which requires to compute the probability of the measurements $p(\mathbf{z}^k|\mathbf{x}^{k-1,(i)})$. To do so, the whole state space at time k should be integrated as :

$$p(\mathbf{z}^k|\mathbf{x}^{k-1,(i)}) = \int p(\mathbf{z}(k)|\mathbf{x}(k))p(\mathbf{x}(k)|\mathbf{x}^{(i)}(k-1))d\mathbf{x}(k). \tag{3.39}$$

In practice, the use of the optimal importance density is difficult to achieve since one has to be able to integrate over the state space and sample particles in $p(\mathbf{z}^k|\mathbf{x}^{k-1,(i)})$. When these conditions cannot be satisfied, a non optimal importance density function has to be chosen.
A popular choice for the importance density function is the transitional prior $p(\mathbf{x}(k)|\mathbf{x}^{(i)}(k-1))$, since in this case the importance weights are simply computed as :

$$w^{k,i} \propto w^{k-1,(i)} p(\mathbf{z}(k)|\mathbf{x}^{(i)}(k)) \tag{3.40}$$

Sampling particles in $p(\mathbf{x}(k)|\mathbf{x}^{(i)}(k-1))$ is generally easier than in $p(\mathbf{x}^k|\mathbf{x}^{k-1},\mathbf{z}^k)$, however this distribution is also more diffuse than $p(\mathbf{x}^k|\mathbf{x}^{k-1},\mathbf{z}^k)$, which makes particles to spread in the state space and may lead to important degeneracy problems.

3.2.6 Density estimation methods summary

We summarize in Table 3.1 the main filtering techniques for density estimation which have been presented above. The study of the characteristics of these methods have highlighted three of them which are of a particular interest for biological particle tracking :

- In the simple linear white and Gaussian case, the KF is the preferred technique since it provides the exact density estimation in the MAP and MSE sense. Moreover, KF enjoys a very low computation cost. In biology, however, the state evolution is often difficult to model with a single KF.
- In the nonlinear/non Gaussian case, the SMC approach is an attractive solution since it relies on very few assumptions. The SMC flexibility makes it therefore feasible in a number of complex cases which are impossible to model with a linear and Gaussian models. The flexibility of the method however comes with the price of a higher cost than KF and its derivatives, and with a non deterministic behavior. Some critical theoretical and computational issues are also to be dealt with such as sample degeneracy and the choice of the importance density.
- The IMM is an appealing solution in between KF and PF since it is able to model complex posterior densities at a relatively low cost when embedding few cheap filters such as KF and its derivatives. When using the IMM the state posterior is approximated with a HMM model which mixes few simple pdfs. This HMM approximation is not feasible in the most

Method	Assumptions	Deterministic	Approximations
KF	linear model with Gaussian Noise and Gaussian state posterior	yes	no approximation, optimal MAP and MSE solution
EKF	Gaussian Noise and Gaussian state posterior	yes	linearization of state functions
UKF		yes	Gaussian approximation of the state posterior
IMM	assumptions used by the embedded state models	yes	HMM state approximation
SMC		no	state posterior approximation with random samples

Tab. 3.1: Main filtering methods for density estimation

general case, but it is appropriate to a number of situations in which the state posterior is multimodal. We will see in Chapter 5 that the motion of biological particles may induce this kind of density.

In summary, each application should lead to a specific estimation solution since the state space definition depends on the different measurement modalities and targets motion. Cheap solutions such as KF are to be favored when they are allowed by the state model since more complex methods may suffer from a number of theoretical issues and induce a higher computational cost. For complex situations, advanced techniques such as SMC or IMM are efficient solutions. In the next section we browse through a number of tracking algorithms, each embedding a different filtering step adapted to the biological system for which they were designed.

3.3 Bayesian tracking in a frame-to-frame approach

In the field of cellular imaging state space filtering techniques were first used to track particles in two dimensions [125] and then in three dimensions [70]. Over recent years these pioneering works have inspired more advanced tracking techniques based on Bayesian concepts [67, 155, 157, 75, 74]. We describe these approaches now, and highlight the main differences between them for the estimation of particle motion and the construction of tracks.

3.3.1 Instantaneous Maximum Likelihood association

To the best of our knowledge, the approach proposed in [125] is the first to introduce the state space model for multiple biological particle tracking. This first work has been then refined over years in a continuous effort to improve the tracking model [70, 72, 71, 65, 66, 69, 68]. The most recent 3-D tracking algorithm [67] directly based on this approach includes many improvements, such as unknown and varying particle motion modeling, which has allowed it to be successfully used in a number of biological applications [3, 33, 6, 116]. We describe here this refined version of the early Bayesian ideas for biological particle tracking.

The tracking procedure proposed in [67] is a frame-to-frame approach relying on maximum likelihood principles, hence we call it call it the *Instantaneous Maximum Likelihood* (IML) method in the remaining of the dissertation. The IML approach deals separately with the spot detection task and the linking of the measurements to form trajectories. The detection of spots is carried out by means of 3-D discrete wavelet transformation of the image stacks and the correlation analysis of wavelet coefficients [129]. The extracted detections Z^l serve as a basis for trajectories construction, as for point correspondence tracking algorithms.

In the state space model, the measurement $\mathbf{z}(k)$ for a given target at time k corresponds to the detections it produces in the stack $I(k)$. The measurement $\mathbf{z}(k)$ is a column vector containing the Cartesian coordinates of the center of mass of the detection : $\tilde{x}(k)$, $\tilde{y}(k)$ and $\tilde{z}(k)$. Moreover, two spot appearance features : the volume of the segmented spot $\tilde{v}(k)$ and its mean intensity $\tilde{i}(k)$, are appended to the vector in order to improve the discriminating power of the posterior pdf of the track state (Eq. 3.18). The measurement $\mathbf{z}(k)$ is therefore written as follows :

$$\mathbf{z}(k) =[\tilde{x}(k), \tilde{y}(k), \tilde{z}(k), \tilde{v}(k), \tilde{i}(k)]^T, \tag{3.41}$$

The IML algorithm sequentially appends detections to existing tracks based on the probability of the detections to originate from each track. To do so, the state of each target has to be estimated from the sequence of noisy measurements it is associated to.

State space filtering with IMM

In [67] the following state definition was proposed for a given particle :

$$\begin{aligned}\mathbf{x}(k) =&[x(k), y(k), z(k), x(k-1), y(k-1), z(k-1)\\ &x(k-2), y(k-2), z(k-2), v(k), i(k)]^T.\end{aligned} \tag{3.42}$$

The vector contains the current Cartesian coordinates of the target, its appearance features, as well as the coordinates of the particle in the two previous frames. In this case, the observation model $\mathbf{H}$ has a simple linear form :

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.43}$$

Moreover, the measurement noise is supposed to be white and Gaussian for each measured feature.

The state transition model has to account for the motion type of particles since the evolution of the particle coordinates between two subsequent frames corresponds to its movement during this time interval. In [67] it was proposed use the IMM filter with several dynamic models to account for the various dynamics of biological objects and the abrupt variations between them over time. Three models of dynamics are defined : random walk (RW), first-order linear extrapolation (FLE), and second-order linear extrapolation (SLE). This choice is supported by the movements of particles wich are often observed in cells [108] : RW models the Brownian motion [134, 182], while FLE and SLE model constant speed and constant acceleration movements, respectively. The mean intensity and volume of spots are assumed to remain approximately constant during the transition between two successive frames, with small random variations described by a Gaussian distribution for each one :

$$p(i(k)|i(k-1)) = \mathcal{N}(i(k); i(k-1), \sigma_i) \tag{3.44}$$

$$p(v(k)|v(k-1)) = \mathcal{N}(v(k); v(k-1), \sigma_v). \tag{3.45}$$

In the following we split the state vector as : $\mathbf{x}(k) = \mathbf{x}^c(k) \cup \mathbf{x}^a(k)$, since kinetic features : $\mathbf{x}^c(k) = [x(k), y(k), z(k)]^T$, and appearance features : $\mathbf{x}^a(k) = [v(k), i(k)]^T$ are assumed to evolve independently in [67].

Random walk (RW)

RW is used to model the Brownian motion [134, 182] for which the pdf of the particle position at time $k+1$ follows a multidimensional normal distribution centered on the previous position. The state transition is thus written as follows :

$$p(\mathbf{x}^c(k+1)|\mathbf{x}^c(k)) = \mathcal{N}(\mathbf{x}^c(k+1); \mathbf{x}^c(k), \mathbf{Q}). \tag{3.46}$$

In this case, the state transition equation can be written linearly, and the corresponding state transition matrix $\mathbf{F}^c_{RW}$ and noise matrix $\mathbf{Q}$ are written as follows :

$$\mathbf{F}^c_{RW} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{Q^c}_{RW} = \begin{bmatrix} \sigma_x^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_y^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_z^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

First-order linear extrapolation (FLE)

FLE is used to model linear movements with constant speed. In this case $\mathrm{d}_{\mathbf{x}}(k)$, the displacement between frames k and $k+1$, is equal to $\mathrm{d}_{\mathbf{x}}(k+1)$, the displacement $\mathrm{d}_{\mathbf{x}}(k)$ between frames $k+1$ and $k+2$. In [67] the equality of displacements between subsequent frames is exploited as follows :

$$\begin{aligned} & \mathrm{d}_{\mathbf{x}}(k+1) = \mathrm{d}_{\mathbf{x}}(k) \\ \Rightarrow & \begin{pmatrix} x(k+1) \\ y(k+1) \\ z(k+1) \end{pmatrix} - \begin{pmatrix} x(k) \\ y(k) \\ z(k) \end{pmatrix} = \begin{pmatrix} x(k) \\ y(k) \\ z(k) \end{pmatrix} - \begin{pmatrix} x(k-1) \\ y(k-1) \\ z(k-1) \end{pmatrix} \\ \Leftrightarrow & \begin{pmatrix} x(k+1) \\ y(k+1) \\ z(k+1) \end{pmatrix} = 2 \begin{pmatrix} x(k) \\ y(k) \\ z(k) \end{pmatrix} - \begin{pmatrix} x(k-1) \\ y(k-1) \\ z(k-1) \end{pmatrix} \end{aligned} \tag{3.47}$$

By using the Equation 3.47, estimating the velocity and the direction of the movement can be avoided in the state transition equation : the velocity and direction are progressively evaluated and updated over time thanks to the previous coordinates. In [67] a linear transition model with the corresponding state transition matrix $\mathbf{F}_{FLE}$ is defined as : small

$$\mathbf{F}^c_{FLE} = \begin{bmatrix} 2 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

by exploiting Equation 3.47. In order to model uncertainty on the directed motion, Gaussian noise independent for each direction of the Cartesian system is assumed, which results in a state noise covariance matrix $\mathbf{Q}^c_{FLE}$ similar to $\mathbf{Q}^c_{RW}$ but with specific parameters. When using the FLE model the state posterior density can be computed as :

$$p(\mathbf{x}^c(k+1)|\mathbf{x}^c(k), \mathbf{x}^c(k-1)) = \mathcal{N}(\mathbf{x}^c(k+1); 2\mathbf{x}^c(k) - \mathbf{x}^c(k-1), \Sigma), \tag{3.48}$$

which involves the combination of independent Gaussian densities (the covariance matrix Σ is diagonal) centered on the predicted state : $2\mathbf{x}^c(k) - \mathbf{x}^c(k-1)$.

Second order Linear extrapolation (SLE)

Curvilinear movements are characterized by a locally almost constant acceleration in the Cartesian system. Hence, in [67] authors have proposed to exploit this property to write a linear constant acceleration model. The idea is to write the acceleration at time k as the difference between the velocity at time k and $k-1$. In a constant acceleration system the following equality can thus be written :

$$\mathrm{d}_{\mathbf{x}}(k+1) - \mathrm{d}_{\mathbf{x}}(k) = \mathrm{d}_{\mathbf{x}}(k) - \mathrm{d}_{\mathbf{x}}(k-1) \tag{3.49}$$

The Equation 3.49 can be developed to express the coordinates of the particle in the frame $k+1$ as a linear function of its coordinates at time k, $k-1$ and $k-2$ as follows :

$$\begin{pmatrix} x(k+1) \\ y(k+1) \\ z(k+1) \end{pmatrix} = 3 \begin{pmatrix} x(k) \\ y(k) \\ z(k) \end{pmatrix} - 3 \begin{pmatrix} x(k-1) \\ y(k-1) \\ z(k-1) \end{pmatrix} + \begin{pmatrix} x(k-2) \\ y(k-2) \\ z(k-2) \end{pmatrix} \tag{3.50}$$

The Equation 3.50 has been used to define a linear state transition equation with the following transition matrix :

$$\mathbf{F}^c_{SLE} = \begin{bmatrix} 3 & 0 & 0 & -3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & -3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & -3 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix},$$

and the state posterior probability can be computed from a multidimensional Gaussian density defined as :

$$\begin{aligned} &p(\mathbf{x}^c(k+1)|\mathbf{x}^c(k), \mathbf{x}^c(k-1), \mathbf{x}^c(k-2)) \\ &= \mathcal{N}(\mathbf{x}^c(k+1); 3\mathbf{x}^c(k) - 3\mathbf{x}^c(k-1) + \mathbf{x}^c(k-2), \Sigma). \end{aligned} \tag{3.51}$$

As shown above, in [67] three linear models have been proposed to account for the various motion types observed in practice. The linear definitions of RW, FLE and SLE allow using a standard KF for each one, hence resulting in a very cheap IMM filter which can be used to model various motions with abrupt changes.

Maximum likelihood association

At a given time k, associations between active tracks and the detections in the current frame are selected following a ML criterion. First the likelihood matrix $\mathbf{L}$ is built as :

$$\mathbf{L} = \begin{bmatrix} & \vdots & \\ \cdots & l_{i,j}^{max} & \cdots \\ & \vdots & \end{bmatrix}, \tag{3.52}$$

by computing for each target i and each detection $\mathbf{z}_j \in Z(k)$ the maximum likelihood $l_{i,j}^{max}$ among the three filters in the IMM :

$$l_{i,j}^{max} = \max_{q \in \{1,2,3\}} p(\mathbf{z}_j(k)|\mathbf{x}_i^{k-1}, \mathbf{z}_{t_i}^{k-1}, r_k = q), \tag{3.53}$$

where $\mathbf{z}_{t_i}^{k-1}$ is the set of measurements assigned to the track i up to time $k-1$, and $r_k = 1$ when the hypothesized motion model at time k is the RW, $r_k = 2$ for the FLE, and $r_k = 3$ for the SLE. Each conditional probability $p(\mathbf{z}_j(k)|\mathbf{x}_i^{k-1}, \mathbf{z}_{t_i}^{k-1}, r_k = q)$ is computed using the KF of the q-th KF.

The assignments are then found by selecting iteratively the element with the greatest value $l_{i,j}^{max}$, deleting then the corresponding row i and line j in $\mathbf{L}$, and repeating the procedure until $\mathbf{L}$ is empty. The complexity of this iterative association selection procedure is very low : once the cost matrix $\mathbf{L}$ is built the computation time is similar to the lNN one. In order to reduce further the association problem complexity and improve the robustness of the algorithm to false detections a search gate discarding unlikely associations is defined. The search gate corresponds to an area containing the measurement at time k with a probability $\alpha = 0.95$. In practice this area is defined around the predicted measurement provided by the IMM of the considered track. Search gates are used to cluster the association problem in a set of smaller association problems which are resolved independently with the iterative ML selection procedure.

Track initiation and termination

The IML has been designed for applications in which the number of targets is unknown and can change over time due to the appearance and disappearance of some particles. Specific rules have therefore been defined to initiate and terminate the tracks if required.

After the association step some detections may remain unassigned, while some tracks did not find any probable measurement. Every unassigned detection is used to initiate a new track to model the possible appearance of a target. By doing so, both true and false are the basis for the creation of a track. In order to discard the resulting false tracks a post-processing step is applied afterwards to discard very short tracks. It is indeed assumed that a false detection cannot find a long sequence of likely detections in the following frames.

When a track does not find any likely detection in a given frame, two cases are considered : (1) the target is missed by the detection procedure because of the high level of noise or its low intensity ; (2) the target does not exist any more in the surveillance volume or has bleached to a low intensity level. In the first case the track should be continued since the target will be present in future frames, while in the second case the track has to be ended. To solve this issue it is proposed in [67] to continue every track without a measurement with some predicted measurements provided by the IMM filter during a fixed number of frames d. An existing target is expected to be detected again within the d following frames, while it is assumed that a track will not find any measurement if the corresponding target does not exist anymore. Hence, every track is ended when no detection is assigned to it during d subsequent frames.

Summary

The IML is the first multiple particle tracking algorithm using extensively the state space model to characterize the tracking problem. As opposed to traditional particle tracking techniques, the detections and the particle states are distinguished to model the measurement process inaccuracy. Moreover a statistical model of the target behavior has been proposed. Thanks to the IMM filtering step various motion types can be modeled, tracking particle with varying motion types is also feasible. In practice the IML works in a very similar way as the lNN do, the distance between detections and tracks is just replaced by the probability that detections originate from tracks. This simplicity results in a computational efficiency which makes the IML feasible for real time applications even if particles are very numerous.

In the IML algorithm no strong assumption is made on the image type, moreover the capability to model different motion types makes it appropriate for the tracking different biological targets. The IML has thus been used in a wealth of tracking applications, such as the tracking of plasmodium sporozoites [3], single locus tracking in live yeast [33], 3D+t tracking of VIH-1 virus complexes during cell infection [6] and 2D+t tracking of bacteria adhering to surface [116]. The IML however suffers form important limitations which make it non feasible in a number of biological applications.

The main weakness of the IML is shared with point-correspondence based tracking techniques : target appearance/disappearance events are not explicitly modeled, and false detections presence is not accounted for in the tracking process. Hence, we have experienced in poor imaging conditions, or when the particle density is high, that the tracking results are significantly degraded because of these unresolved issues. As an example, when the number of false detections is high, it is probable that tracks find some false detections in their search gate. Thus, wrong associations with false detections can occur, and the simple criterion for track termination is not suited since a track which should be ended can easily continue through a series of false detections.

3.3.2 Multiple object tracking with particle filtering state estimation

In [155] it is proposed to avoid the particle detection step by changing the measurement equation to account for particles intensity in images : the state of a target $\mathbf{x}(k)$ at time k is inferred from previous states and the current image data $\mathbf{z}(k)$ without a strict association to any detection. Tracking without a detection step avoids some important issues such as missing and spurious detections which are encountered in standard point correspondence tracking methods. However, by doing so the measurement equation becomes non linear, which requires the use of non linear estimation techniques such as SMC methods, called here Particle Filters (PF), which have been successfully used for target tracking in other fields [86, 177, 104].

The proposed approach is to estimate first the posterior pdf $p(\mathbf{x}_i(k)|\mathbf{z}^k_{t_i})$ for each target i, with $i = 1..n$, at each time $k = 1..l$ in an iterative way. Once this state posterior is estimated the target state sequence can be inferred on the basis of the expectation, MAP estimation, and MMSE estimation. In the case of a single target the estimation of the distribution $p(\mathbf{x}(k)|\mathbf{z}^k)$ can be achieved with standard PF algorithms, however the task is complicated by the presence of an unknown number of other targets, the complexity of target dynamics and the random target appearance and disappearance.

Some solutions to biological particle tracking issues with PF have been proposed in [155] and then in [157]. We present briefly the particle tracking procedure proposed in [157] which includes some refinements over the method in [155], making it feasible in a broader range of applications.

State and measurement definition for a single target

The state of a target is defined in [157] as a vector :

$$\begin{aligned}\mathbf{x}(k) &= [x(k), \dot{x}(k), y(k), \dot{y}(k), \sigma_{max}(k), \sigma_{min}(k), I(k)]^T, & (3.54)\\ &= [\mathbf{r}(k), \mathbf{v}(k), \mathbf{s}(k), I(k)]^T, & (3.55)\end{aligned}$$

where the target position $\mathbf{r}(k) = [x(k), y(k)]^T$ and its velocity $\mathbf{v}(k) = [\dot{x}(k), \dot{y}(k)]^T$ are accounted for. Non kinetic information, such as the object shape feature vector $\mathbf{s}(k) = [\sigma_{max}(k), \sigma_{min}(k)]^T$ and its intensity $I(k)$, is also included in the state model. Assuming that the changes in the position, intensity, and shape parameters are independent, the conditional pdf $p(\mathbf{x}(k)|\mathbf{x}(k+1))$ can be factorized as :

$$p(\mathbf{x}(k+1)|\mathbf{x}(k)) = p(\mathbf{y}(k+1)|\mathbf{y}(k)) \cdot p(\mathbf{s}(k+1)|\mathbf{s}(k)) \cdot p(I(k+1)|I(k)), \tag{3.56}$$

where $\mathbf{y}(k) = [\mathbf{r}(k), \mathbf{v}(k)]^T$ contains the kinetic features at time k. Authors propose to deal with multiple dynamics by using a jump Markov model, as in the IMM, between two modes of motion : random walk motion and nearly constant speed motion with small accelerations. The probability to switch between the two modes is defined by a two states Markov chain with predefined transition probabilities.

The random walk model ($i = 1$) is the same as in [67], so a transition matrix $\mathbf{F}_1$ and a noise matrix $\mathbf{Q}_1$ are defined in a similar way as $\mathbf{F}_{RW}$ and $\mathbf{Q}_{RW}$. On the other side, the constant velocity model ($i = 2$) slightly differs : the state transition is still linear as in Equation 3.8, but the transition and noise matrices for $\mathbf{y}$ are written as in [14] :

$$\mathbf{F}_2 = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{Q}_2 = q_{1,2} \begin{bmatrix} T^3/3 & T^2/2 & 0 & 0 \\ T^2/2 & T & 0 & 0 \\ 0 & 0 & T^3/3 & T^2/2 \\ 0 & 0 & T^2/2 & T \end{bmatrix},$$

where T is the time sampling interval. The definition of the state and the transition matrix explicitly use the velocity $\mathbf{v}(k)$, while in [67] the constant velocity property is implicitly accounted for by modeling a constant displacement between frames (Eq. 3.47). During the motion models initialization some coefficients, such as $q_{1,2}$ and the parameter of diffusion in the random walk model of motion, need to be tuned experimentally to the application.

Changes in shape features are modeled with a Gaussian law : $p(\mathbf{s}(k+1)|\mathbf{s}(k)) = \mathcal{N}(\mathbf{s}(k+1); \mathbf{s}(k), Tq_2\mathbf{I})$, where the setting of the q_2 parameter accounts for the magnitude of the random fluctuations.

The pdf of the intensity transition $p(I(k+1)|p(I(k)))$ needs to reflect the process of bleaching process. To do so, authors assume an exponential decay over time of the intensity. The corresponding decay parameter α is estimated beforehand by fitting it to evolution over time of the background intensity. A simple first order Gauss-Markov process assumption is used here for the intensity transition prior : $p(I(k+1)|I(k)) = \mathcal{N}(I(k+1); (1-\alpha)I(k)), q_3T)$, where the parameter q_3 is fixed beforehand, according to the level of intensity random fluctuations.

Observation model

The measurements $\mathbf{z}^l$ is chosen directly as being the set of images over time I^l in [157]. The image formation process is a standard additive model where the intensity profile of each target is superimposed to a constant background. A Gaussian PSF and spherical nano-metric scale objects are both assumed, which leads to approximate particle intensity profile as being a multidimensional Gaussian function. The intensity of elongated objects are also modeled by utilizing the velocity components from $\mathbf{x}(k)$ as a parameter in the Gaussian function when it is assumed that objects move in the same direction as their elongation. For a single target at position $\mathbf{r}(k)$, the intensity in pixel (i, j) is modeled as the function h_k of the target state :

$$h_k(i,j;\mathbf{x}(k)) = I(k)\exp(-\frac{1}{2}\mathbf{m}_k^T\mathbf{R}^T\Sigma_k^{-1}\mathbf{R}\mathbf{m}_k) + b_k(i,j). \tag{3.57}$$

In the function $h_k(i,j;\mathbf{x}(k))$ the rotation matrix $\mathbf{R}$ is linked to the target velocity at time k, the vector $\mathbf{m}(k)$ accounts for the distance to the particle centroid, while the diagonal matrix $\Sigma_k = diag[\sigma^2_{max,k}, \sigma^2_{min,k}]$ models the blurring process in the direction of the object elongation and in its orthogonal direction. The background b_k is assumed constant and estimated from the images by fitting a mixture of Gaussian to the intensity histogram.
In the Equation 3.57 the dependency of the observation to the state components such as the position, or the velocity, is highly non linear. Hence estimation techniques such as the KF can not be used anymore, which has led the authors to use the PF approach.

Multiple target state estimation

Tracking multiple targets in the PF framework can be done by combining each target state vector to form a single vector : $\mathbf{x}(k) = \cup_{i=1}^{n}\mathbf{x}_i(k)$. While being the most valid approach, doing so would result in an exponential explosion of the computational demand due to the large number of particles required to get accurate approximations in high dimension. The approach adopted in [157] is therefore to use a joint PF only when targets are close to each other, and to run an independent PF for targets which are far from other objects.
The posterior density $p(\mathbf{x}(k)|\mathbf{z}^k)$ is generally multimodal for two main reasons : (1) an object state is often ambiguous due to inaccurate measurements ; (2) when particles are close to each other the multimodality comes from the different combinations between objects and measurements, as for the point correspondence problem. However, maintaining the multimodality in $p(\mathbf{x}(k)|\mathbf{z}^k)$ is an issue for which standard PF methods are poor at. That's why it has been proposed in [157] to explicitly model the multimodality in the posterior distribution with a mixture of M individual non-parametric distributions :

$$p(\mathbf{x}(k)|\mathbf{z}^k) = \sum_{m=1}^{M} \pi_{m,k} p_m(\mathbf{x}(k)|\mathbf{z}^k), \tag{3.58}$$

as done in [176]. The set of coefficients $\{\pi_{m,k}\}_{m=1..M}$ corresponds to the weights of the M component mixture with the normalization constraint $\sum_{m=1..M}\pi_{m,k} = 1$. Component weights are updated on line following a standard Bayesian sequential estimation procedure [176] and when the posterior density $p(\mathbf{x}(k)|\mathbf{z}^k)$ becomes too diffuse, a re-clustering procedure [176], based on the K-means algorithm, is used to build M new components from the existing set of particles. By doing so, a limited set of significant tracking hypotheses is extracted from the diffuse particle distribution.

Filtering and data dependent sampling

In [157], authors choose to apply the Marginal Particle Filtering (MPF) [105] instead of using a re-clustering procedure to deal with the particle degeneracy issue. Indeed, in the standard PF approach described in Section 3.2.5, particles are sequentially drawn from the pdf $p(\mathbf{x}^k|\mathbf{z}^k)$ or from the importance distribution $q(\mathbf{x}^k|\mathbf{z}^k)$. Doing so results in a degeneracy of the particles because of the monotonic increase of the dimensionality of the density to draw samples in. On the contrary, in the MPF particles are drawn from the marginal distribution $p(\mathbf{x}(k)|\mathbf{z}^k)$, resulting in a lower variance of the importance weights. In order to take significant benefits from the MPF the importance sampling function has to depend on the measurement $\mathbf{z}(k)$: using the transition prior $p(\mathbf{x}(k)|\mathbf{x}(k-1))$ would result in similar issues as with the standard PF.

In [157] it is proposed to draw particles from a data-dependent sampling function in order to benefit from the low degeneracy property of the MFP. The sampling function is a mixture of a state prior $p(\mathbf{x}(k)|\mathbf{x}(k-1), i)$ accounting for the mode of motion i, and a data-dependent prior distribution as follows :

$$\begin{aligned} q(\mathbf{x}(k)|\mathbf{x}(k-1), \mathbf{z}(k), i) =& \gamma p(\mathbf{x}(k)|\mathbf{x}(k-1), i) \\ &+ (1-\gamma)\tilde{q}(\mathbf{y}(k)|\hat{\mathbf{y}}(k-1), \mathbf{z}(k), i) \\ &\cdot p(\mathbf{s}(k)|\hat{\mathbf{s}}(k-1))p(I(k)|\hat{I}(k-1)), \end{aligned} \tag{3.59}$$

where $\tilde{q}(\mathbf{y}(k)|\hat{\mathbf{y}}(k-1), \mathbf{z}(k), i)$ depends on the mode of motion i. In practice, the procedure to draw N_p particles from the distribution 3.59 with a mixture parameter $0 \leq \gamma \leq 1$ is the following : γN_p particles are drawn from the first distribution of the mixture, and then $(1-\gamma)N_p$ particles are drawn from the second distribution.

In order to reduce the complexity of the PF approach authors propose to factorize the joint state variables distribution as follows :

$$p(\mathbf{y}(k), \mathbf{s}(k), I(k)|\mathbf{z}^k) = p(I(k)|\mathbf{y}(k), \mathbf{s}(k), \mathbf{z}^k)p(\mathbf{y}(k), \mathbf{s}(k)|\mathbf{z}^k), \tag{3.60}$$

which is known as Rao-Blackwellization [54]. The transition and observation models for the intensity $I(k)$ are Gaussian and linear, hence the pdf $p(I(k)|\mathbf{y}(k), \mathbf{s}(k), \mathbf{z}^k)$ is estimated with a standard KF for each sample path. The MPF is used in this case to estimate $p(\mathbf{y}(k), \mathbf{s}(k)|\mathbf{z}^k)$ instead of $p(\mathbf{y}(k), \mathbf{s}(k), I(k)|\mathbf{z}^k)$ leading to a reduction of the dimension. Marginalization and Rao-Blackwellization are therefore expected to alleviate the particle degeneracy issue.

Track management

Tracking particles in microscopy image sequences requires to deal with events such as target appearance and disappearance, which force the algorithm to estimate the number of targets, end some tracks, and create new ones in each frame. In [157], detections provided by the algorithm based on the h-dome transform and sample clustering are used to initialize tracks in the first frame. The detection procedure is then applied to the whole sequence of images. In each frame, the number of tracks is sequentially compared to the number of detections : it the number of detections exceeds the number of tracks, then new tracks are initialized from the exceeding positions.

Tracks termination is not based on the presence of detections since the tracking algorithm can miss some targets due to their low intensity. Instead, the track termination decision is based on the analysis of the unnormalized importance weights (Eq. 3.34). It is checked whether the average of the unnormalized likelihood values $p(\mathbf{z}(k)|\mathbf{x}_i(k))$ of all particles have dropped below a threshold π_d, which corresponds to having a measurement from the background. In this case it is considered that the tracked object has disappeared and the corresponding tracks is ended.

Summary

PF has been shown in [155] and in [157] to be an elegant framework for tracking particles in microscopy images. Using nonlinear observation models allows modeling the image information and including it directly in the state space. Doing so avoids the inherent problems of using a fixed detection set such as missed detections and false detections. The methods is also able to track non isotropic objects and was recently refined to include multiple dynamics models. The PF based tracking algorithm has been tested with a number of biological data such as 2D images of microtubules, Golgi units, and androgen receptors [155, 157]. A similar approach has been recently adopted to track HIV-1 particles in 2D image sequences [75, 74].

While being theoretically appealing, using PF techniques introduces some major issues avoided by point correspondence based methods. A number of issues are inherent to the use of PF technique, such as the degeneracy of particles problem or the choice of the sampling distribution. Moreover, PF is an estimation technique which cannot deal in a straightforward way with events such as track creation/deletion. Thus, in practice, the use of a detection step is still required. For each of these issues a specific technique has to be used, which increases the computational cost and the number of parameters to tune. Tracking very high density of particles also results in a dramatic increase of the computational demand since independent PF cannot be use anymore for each particle. In this case, avoiding tracks coalescence is also an important issue since the overlaps between different particles are not explicitly prevented.

3.4 Conclusion

We have provided in this chapter an overview of probabilistic methods for particle tracking. The main interest of such approaches is the design of tracking score functions which correspond to a statistical behavior of particles. The statistical models are able to account for the different motion types of particles, as well as other features such as intensity fluctuations and the temporary disappearance of blinking targets. The state space model offers a flexible means of including the different features in a common framework. By doing so, a number of powerful estimation techniques can be used to provide accurate estimates of particle states and various pdfs which can be used for particle tracking. Using a statistical description of the information used for tracking particles solve the issue of information fusion from which methods presented in the Chapter 2 suffer. Moreover, the information quality is addressed in this framework by the use of proper pdfs for the different features. As an example the inaccuracy of the measurements is accounted for, as well as the features fluctuations and the randomness of the particle movements.

The benefits of using statistical concepts for particle tracking are numerous, however the techniques based on these ideas are still unable to process a number of image sequences. In particular, poor imaging conditions and the particle appearance/disappearance events are generally not handled properly. The state-of-the-art algorithms have indeed adopted a very local approach in which the tracking problem is centered on the current frame for a set of given tracks. The presence of false detections, the appearance of new targets and the disappearance of some particles are not accounted for when extending tracks in the current frame, but treated separately with a set of simple rules. Moreover, the numerous sources of uncertainty when tracking dense particles in poor imaging conditions may result in very diffuse statistical models which provide barely significant scores for different tracking solutions. This issue is emphasized by the limited use of the temporal information which is done by the standard algorithms : only the information prior to the current frame is accounted for, while the future frames are ignored.

Part II

Advances in probabilistic particle tracking

Chapter 4

Improved particle detection in complex backgrounds with blind source separation

4.1 Detecting and tracking spots in complex backgrounds

As shown in Chapter 1, particle localization procedures are generally based on the detection of a local intensity increase. Such approaches rely on an underlying simplistic model which assumes that the image is the addition of the targets intensities to a locally constant background and a random acquisition noise (Eq. 1.7). The constant background assumption is however rarely valid when imaging biological systems in fluorescence microscopy. Indeed, both the intracellular and extracellular areas are very crowded environments and the labeling is generally not strictly specific to the targets. Hence structures such as membranes or filaments emit photons which create a structured dynamic background.

Figure 4.1 shows an image of epidermal root cells of Arabidopsis plants. The seedlings expressed the Green Fluorescent Protein (GFP) fused to the KORRIGAN1 (KOR1) protein and 20 stacks of 11 slices each were acquired with a DSCM. KOR1 is found under two physical forms : aggregated in small compartments and fused to cell walls. In this particular application we are interested in the trajectories of KOR1 compartments only and not in particles fused to walls which are also visible in images. Non desirable protein particles however produce a high intensity and structured background which is not accounted for by the flat background image model (Eq. 1.7). Hence using spot enhancement techniques based on this model does not provide satisfactory results. We have for instance experienced a number of standard spot enhancement techniques presented in Chapter 1 with the plant cells images : top-hat filtering, grayscale top-hat opening, wavelet decomposition and wavelet scales product. We show in Figure 4.2 the enhancement of spots in a small crop of an image of KOR1 with these methods. It is remarkable that whatever the used technique walls are still visible with an intensity that is above the intensity level of several small particles. The enhancement of walls is due to its close similarity to a particle at a very local level and its high intensity. Particle detection methods based on spot shape only cannot thus avoid detecting walls.

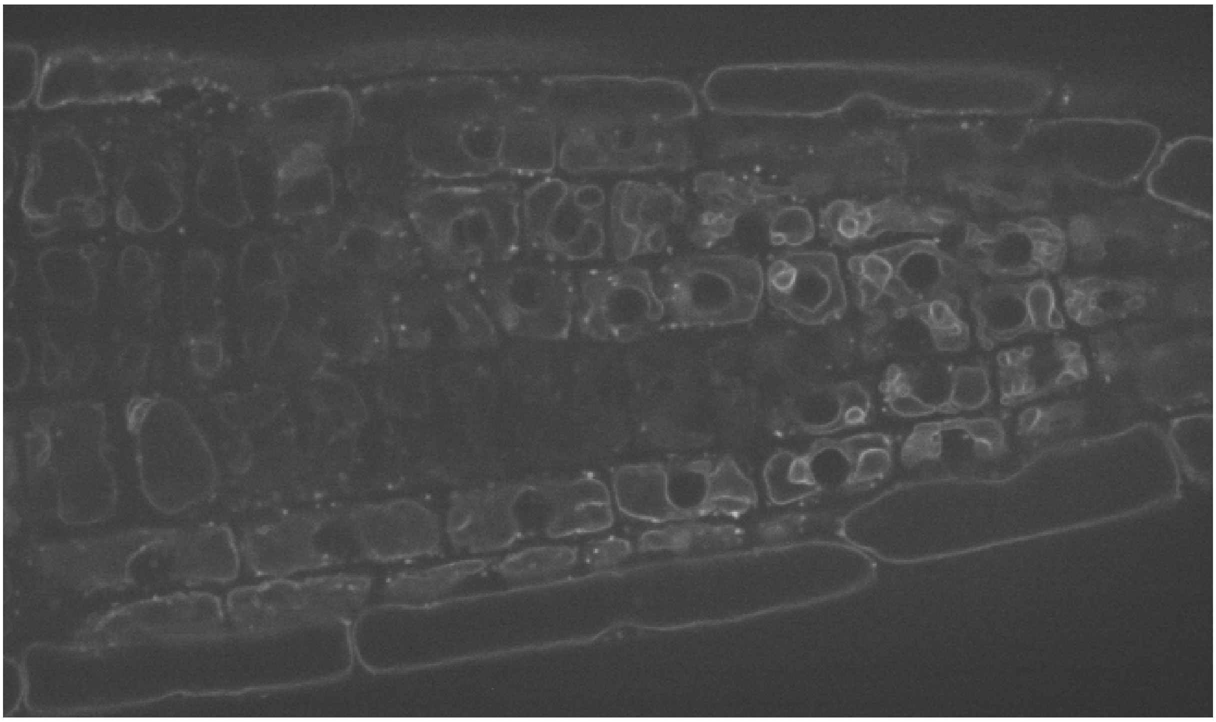


FIG. 4.1: The GFP KOR1 protein labeling vacuole membranes and small compartments in 3D DSCM images.

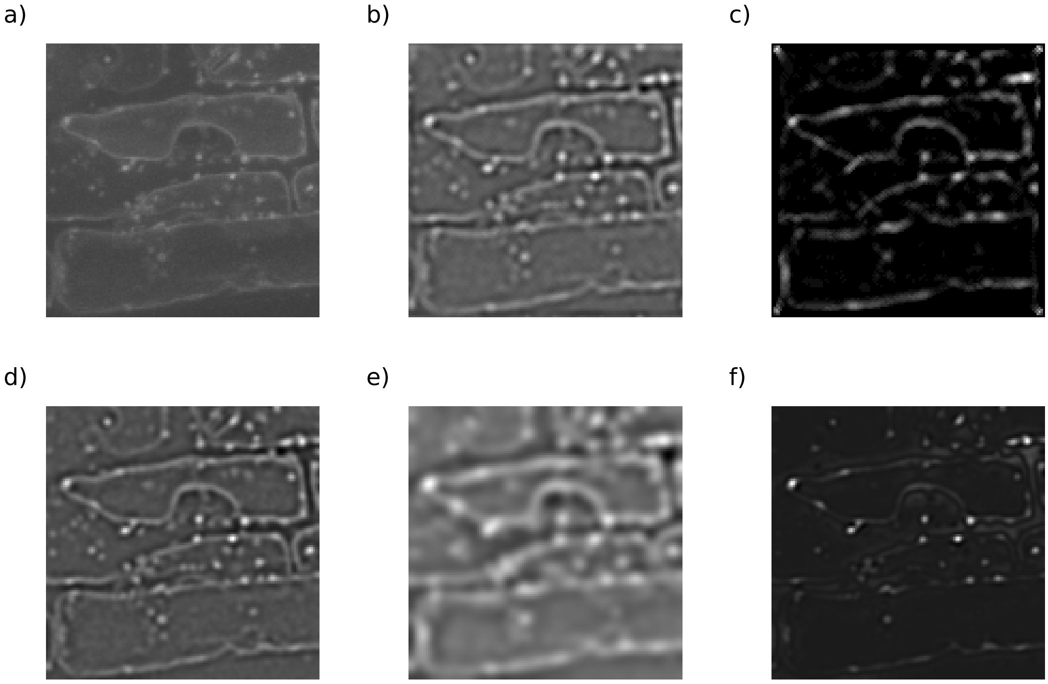


FIG. 4.2: Spot enhancement a in plant cells image. a) original fluorescence image. b) top-hat filtering of. c) grayscale top-hat opening, d) scale 2 of the B3-UDWT. e) scale 3 of the B3-UDWT. f) product of scales 2 and 3 of the B3-UDWT.

Structures that are not accounted for in simple image models produce some photons that can be misinterpreted as a signal of interest by the detection procedure, particle tracking is thus significantly complicated : numerous spurious detections result from the background intensity hence leading to false associations of tracks and the creation of non existent trajectories. Many particles can also be missed due to their aggregation with background structures in images. Moreover, the erroneous image model invalidates the use of tracking method that takes benefits of the image information such as the algorithm proposed in [39] which models both the motion of particles and their intensity profile. We present in this chapter a solution to this issue : we explicitly model background structures, hence leading to an accurate particle detection procedure despite a complex structured background. The augmented image model and its estimation procedure we propose also allow using appearance based tracking procedures.

4.2 Mixed image model and blind source separation

In Chapter 1 we have presented a standard MPG image model which assumes that the Poisson random term is mainly due to three sources : photons counting from the objects in the focal plane, out of focus objects, autofluorescence and the dark noise. Instead of considering that objects in the focal plane are of interest we propose here to distinguish two sources of photons among in focus objects : (1) particles of interest, (2) other background structures. We consequently express the mean intensity of objects in the focal plane λ_i as the sum of the expected background intensity λ_b and the expected target intensity λ_t : $\lambda_i = \lambda_b + \lambda_t$. This model leads us to define a simplified Gaussian model as in Equation 1.7 :

$$I(x, y, z) = T(x, y, z) + B(x, y, z) + W(x, y, z). \tag{4.1}$$

where $T(x, y, z)$ is the intensity value produced by the targets, $B(x, y, z)$ is the background intensity and $W(x, y, z)$ a random noise which is generally assumed white and Gaussian with experimentally estimated moments $(\mu_n(x, y, z), \sigma_n(x, y, z))$.
In practice we measure intensity values in the whole image $\{I(x, y, z)\}_{(x,y,z)\in I}$ but we do not have access to the different terms of the mixture model defined by Equation 4.1. Deciphering a mixture of sources from noisy measurements is similar to a well known inverse problem coined blind source separation (BSS)(see [43] for a review). We thus investigate next the use of BSS techniques for estimating the particle intensity $T(x, y, z)$ in each voxel.

4.2.1 Independent Component Analysis

One of the most widely used BSS technique is the Independent Component Analysis (ICA) [46, 36, 88]. ICA aims at recovering unobserved signals from several mixtures observed with several sensors by applying statistical principles. A simple model assumes the existence of n independent signals $\mathbf{s}(t) \triangleq [s_1(t), \ldots, s_n(t)]^T$ and the observation of the same number of mixtures $\mathbf{y}(t) \triangleq [y_1(t), \ldots, y_n(t)]^T$. The mixtures are obtained as instantaneous linear combinations of the sources :

$$\mathbf{y}(t) = \mathbf{A}\mathbf{s}(t), \tag{4.2}$$

where $\mathbf{A}$, coined the *mixture matrix*, is a $n \times n$ matrix containing the mixture coefficients. The matrix $\mathbf{A}$ is however not known *a priori* : coefficients reflect the physical properties of the emission, transmission and acquisition of the system, such as the spatial repartition of the sensors in the case of a grid of n radar sensors acquiring simultaneously a signal from the same sources.

Modeling the physics of the systems with a high enough accuracy is however a very difficult task due to the high number of free parameters that are involved. The main challenge of BSS is thus not inverting the mixture itself (Eq. 4.2) but to estimate the mixture matrix $\mathbf{A}$ on the fly. The key idea of BSS methods is to avoid a physical modeling of the acquisition process by taking full advantage of the *prior* statistical independence of the sources. Such methods are therefore qualified as *blind* techniques since no prior information on $\mathbf{A}$ is exploited.
ICA technique extracts sources by optimizing a cost function which maximizes sources independence while accounting for some prior information on their statistical distribution. The key idea of the ICA is to take benefits of the diversity of information provided by the sensors to separate the independent sources. Indeed, when multiple sensors are used each one has a different 'point of view' on the 'scene' : each sensor provides a different mixture of sources. The ICA is therefore not suited to our case : each fluorescence images is a view of a scene that is continuously changing and a single sensor is generally used : the microscope camera. Since we cannot exploit diversity we investigate next an approach based on a different criterion to separate sources in a single image.

4.2.2 Morphological Component Analysis

Recently a new source separation technique called Morphological Component Analysis (MCA) has emerged [164]. The algorithm is capable of separating several sources in a single noisy mixture by exploiting some prior morphological differences between them. In the MCA framework the observed signal $\mathbf{Y}$ is a n-dimensional vector assumed to be the sum of p independent sources $\{\mathbf{s}_i\}_{i=1..p}$:

$$\mathbf{Y} = \sum_{i=1..p} \mathbf{s}_i + \mathbf{N}, \tag{4.3}$$

where $\mathbf{N}$ is a Gaussian vector with a diagonal covariance matrix modeling the white and Gaussian acquisition noise $\mathcal{N}$. Source separation consists in the estimation of the vectors $\{\mathbf{s}_i\}_{i=1..p}$ based on the measurement of the single n-dimensional vector $\mathbf{Y}$. To do so the MCA relies on the projection of $\mathbf{Y}$ in a mathematical space in which the sources and the noise are easily separable. Using a linear transformation, the signal $\mathbf{Y}$ is decomposed as :

$$\mathbf{Y} = \mathbf{D}\alpha, \tag{4.4}$$

where α is the coefficient vector of $\mathbf{Y}$ in the representation dictionary D. The matrix $\mathbf{D}$ is a linear reconstruction operator of D. The MCA principle is to take benefits of the structure of D to easily separate the sources and the noise. Indeed, the dictionary D is built as the concatenation of a set of p representation dictionaries :

$$\mathbf{D} = [\mathbf{D}_1, \dots, \mathbf{D}_p], \tag{4.5}$$

such that each source is assumed to be represented in a very sparse way in a given waveform dictionary and in a very non-sparse manner in every other. A signal s_i is said to be sparse in a waveform dictionary D_j if it can be represented from very few dictionary elements : $\mathbf{s}_i = \mathbf{D}_j\alpha_{ij}$, where most of the coefficients α_{ij} are zero ($||\alpha_{ij}||_0$ small). Hence, if the construction of $\mathbf{D}$ is appropriate the following property is true :

$$\forall i = 1..p,\ \forall j = 1..p,\ s.t.\ j \neq i :\ ||\alpha_{ii}||_0 \ll ||\alpha_{ij}||_0, \tag{4.6}$$

with an appropriate indexing of the dictionaries. When finding a sparse representation α of $\mathbf{Y}$ in D :

$$Y = \mathbf{D}\alpha \text{ with } ||\alpha||_0 \text{ small,} \tag{4.7}$$

each source $\mathbf{s}_i$ is thus assumed to project onto the sub-dictionary D_i only since projecting $\mathbf{s}_i$ onto other parts of D would produce a much larger number of non-zero coefficients. The key idea of the MCA is therefore to find a sparse representation of the signal $\mathbf{Y}$ in the dictionary D, and then to reconstruct an estimation $\hat{\mathbf{s}}_i$ of each source $\mathbf{s}_i$ from the subset of coefficients α_i corresponding to the i^{th} dictionary as :

$$\hat{\mathbf{s}}_i = \mathbf{D}_i\alpha_i. \tag{4.8}$$

The dictionary D is highly overcomplete since it is made of the union of p possibly overcomplete dictionaries, so in practice the matrix $\mathbf{D}$ is highly non-square and thus non-invertible. The possible projections of the data onto the dictionary are therefore multiple and a specific search procedure has to be designed to find a sparse representation α. Moreover, the observed signal $\mathbf{Y}$ is corrupted by the noise N which should be accounted for. In [164] the following score function is proposed :

$$\alpha^\star = \min_\alpha ||\alpha||_0 + \lambda||Y - \mathbf{D}\alpha||_2^2 \tag{4.9}$$

for retrieving the optimal representation $\alpha^\star$ of the image $\mathbf{Y}$ in $\mathbf{D}$. In Equation 4.9 the first term of the sum reflects a sparsity constraint while the second term is the squared reconstruction error and introduces a data-driven constraint. The trade-off between the sparsity and the error of representation is tuned by the parameter λ which is generally fixed as a function of the noise standard deviation σ_N :

$$\lambda = \frac{1}{k\sigma_n}, \tag{4.10}$$

where $k > 0$ is a free parameter. By doing so the level of noise is accounted for :

1. When the noise level is low λ is high, so the fidelity of the representation α is enforced.
2. A high level of noise relaxes the data-driven constraint whereby the poor reliability of the image is accounted for.

The optimization algorithm proposed in [164] consists in an iterative alternate projection and thresholding scheme which is detailed in Algorithm 1. The procedure relies on the sequential update of each source which is based on a thresholding of the representation when other sources are assumed fixed. In practice for a source $\mathbf{s}_i$ the projection of the residual between the image and other sources is computed and projected onto the dictionary D_i thanks to a forward transform $\mathbf{D}_i'$. This representation is then δ hard-thresholded in order to maximize the MCA objective function (the soft-thresholding operator can be used when the sparsity term is expressed as $||\alpha||_1$). This procedure is alternated between the different sources and with different thresholding settings : the threshold δ decreases as more iterations are computed, so only very high coefficients are first captured in α, and then more coefficients are progressively added until the final threshold $\delta_f = 1/\lambda$ is reached. There exists mainly three evolution strategies for the threshold δ : linear and exponential decrease [164], and Mean-of-Max [22] which requires less iterations for the same quality of separation but with an increased cost for each iteration.

The MCA technique has known important developments over the past few years : it has been

Algorithm 1 The standard MCA algorithm.

```
Require: Initialize $L_{max}$ the number of iterations and the final threshold $\delta^f$
  for $iter = 1$ to $L_{max}$ do
    Update the threshold $\delta$.
    for $i = 1$ to $p$ do
      {The coefficients $\hat{\alpha}_i$ are updated assuming that other sources $s_k$, $k \neq i$ are fixed.}
      Compute the residual $\mathbf{R}_i = \mathbf{Y} - \sum_{j=1,j\neq i}^{p} \mathbf{D}_j \hat{\alpha}_j$.
      Compute the projection onto $\mathbf{D}_i : \alpha_i = \mathbf{R}_i \mathbf{D}'_i$.
      Threshold the coefficients $\alpha_i$ with the $\delta$ threshold to obtain $\hat{\alpha}_i$.
      Reconstruct an estimation of $s_i$ as : $\hat{\mathbf{s}}_i = \mathbf{D}_i \hat{\alpha}_i$
    end for
  end for
  return $\hat{\mathbf{s}} = [\hat{\mathbf{s}}_1, \ldots, \hat{\mathbf{s}}_p]$
```

extended to multi-channels data in [23, 167], a link with ICA has been established in [167], it has been combined with the Total Variation regularization for image compression via separation [165] and with dictionary learning techniques [130, 131]. We refer to [24] and [61] for a comprehensive review of the MCA methodological links and extensions. MCA has thus been applied in several different cases : textures and natural scenes separation in images [164], inpainting in natural images [60], color images restoration [23], spherical data maps processing in astrophysics thanks to wavelet, curvelet and ridgelet transforms on the sphere [1], cosmic microwave background analysis [24] and audio signals separation.

The MCA is an appealing approach for background and particles separation in fluorescence images since these sources have different morphologies and since it does not require multiple sensors or samples as the ICA. The use of the MCA for 3D biological image sequences processing however requires some adaptations to the specificity of the data which we present next. We also introduce in Section 4.4.1 a novel detection procedure based on the MCA technique which allows us to localize reliably particles in complex structured backgrounds and to track them over time.

4.3 MCA for fluorescence image analysis

4.3.1 Representation dictionaries for 2D and 3D microscopy images

The main requirement for using the MCA source separation techniques is the definition of a dictionary set $\{D_i\}_{i=1..p}$ such that each source is sparse in an unique and different dictionary as pointed out by Equation 4.6. To do so sources must have sufficiently different morphologies and each dictionary must be specific to one of these signal morphologies. In fluorescence images small particles appear as Gaussian spots due to the blurring of the acquisition device which can be viewed as a convolution with a Gaussian function : the microscope PSF (see Chapter 1). On the other hand, the background often originates from complex curvilinear structures such as cell and nucleus membranes, actin filaments, microtubules and flagella. For instance KOR1 protein fused to cell walls produce a complex network of curvilinear structures in plant cells images (Fig. 4.1).

The shape differences of particles and other structures in fluorescence images lead us to consider them as two sources in the MCA framework. We represent them as column vectors :

$$\mathbf{s}_1 = [T(1,1,1), \ldots, T(l_x, l_y, l_z)]^T, \tag{4.11}$$

$$\mathbf{s}_2 = [B(1,1,1), \ldots, B(l_x, l_y, l_z)]^T. \tag{4.12}$$

where l_x, l_y and l_z are the image stack width, height and depth, respectively. The two column vectors $\mathbf{s_1}$ and $\mathbf{s}_2$ are therefore n-dimensional with $n = l_x \times l_y \times l_z$. We investigate next the choice of the dictionaries D_1 and D_2 which must represent sparsely one source only and be fast to compute since hundreds of forward and backward transforms are calculated by the MCA scheme. A low redundancy is also desirable since the computational and memory requirements grows with the number of coefficients to process. An additional requirement is the facility of implementation since we aim at implementing a java version of the transforms in order to provide the MCA algorithm to biologists via the Quia sofware developed in the Unité d'Analyse Quantitative d'Images at Institut Pasteur.

Representing particles with the B3-UDWT

As pointed out in Chapter 1 small particles appear as nearly Gaussian spots in fluorescence images since the point spread function of the acquisition system has a multidimensional Gaussian like shape [188]. The B3-UDWT image decomposition is thus particularly well suited for the sparse representation of fluorescent particles in the MCA framework :

- The wavelet kernel is an approximation of a LoG filter, hence representing Gaussian like spots with few high coefficients (see Figures 1.5 and 4.2).
- The wavelet function flattens slowly varying backgrounds. The background therefore does not produce significant coefficients except in the approximation image.
- The redundancy of the transform is moderate : $n \times J$ coefficients are produced for a decomposition in J scales.
- The forward transform is very fast due to the separability of the transform along the x-axis, y-axis and z-axis. Also the one-dimensional transform involves successive convolutions with a small kernel (5 non-zero coefficients) which are thus very fast. The backward transform is extremely fast since it consists in a coefficient by coefficient sum across scales. The computational and memory requirements are therefore very low and scale linearly with n for both the forward and backward B3-UDWT.

The B3-UDWT is able to represent sparsely spots in images, however, as shown in Figure 4.2 it also represents linear structures with a moderate number of coefficients since a curve can be seen locally as a combination of points. We have therefore to select a second dictionary D_1 that is able to represent curvilinear structures structures with a very low number of coefficients to avoid the curves to be misinterpreted as spots by the MCA.

Representing the curvilinear background with the Curvelet transform

The dictionary D_2 representing the background must be selected according to its ability to represent sparsely curvilinear structures and not small spots. Background structures can consist in a very complex arrangement of elements of various length and width, hence a multiscale representation property is highly desirable.

The 2D curvelet transform (CVT1) [34, 53, 163, 164] allows the directional analysis of an image in different scales by decomposing an images as a combination of segments of various length and width. The CVT1 dictionary thus provides an optimally sparse representations of objects which display smoothness except for discontinuities along a general curve with bounded curvature [57].

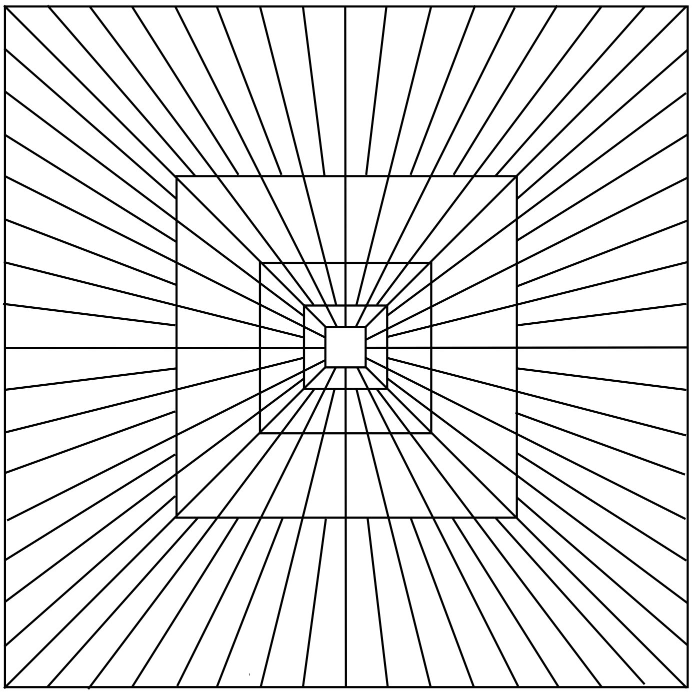


Fig. 4.3: Fourier domain coronization roughly achieved by the CVT2 with four scales and eight angles at the coarsest scale.

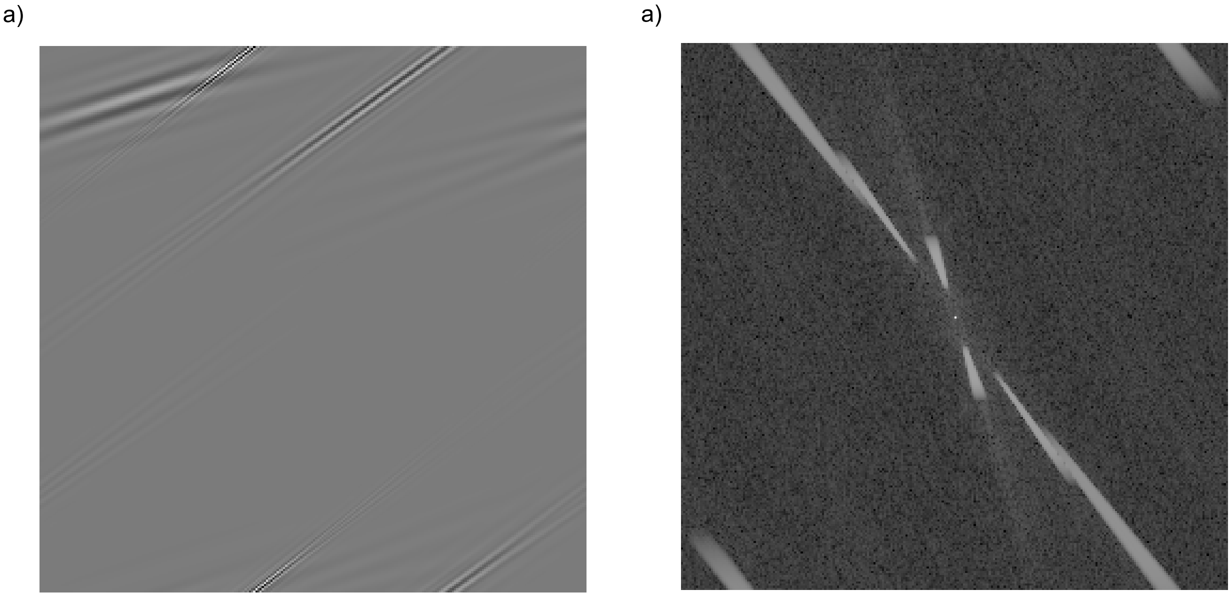


Fig. 4.4: a) Image reconstruction of three curvelet coefficients. b) The discrete FFT of the image a).

In practice the image decomposition is achieved in three steps : wavelet analysis, overlapping block partitioning and discrete ridgelet analysis [35]. The reconstruction is exact due to the invertibility for each step of the decomposition scheme. The redundancy of the representation is however very high : $16J + 1$ where J is the number of wavelet scales. Indeed the redundancy is $4J + 1$ for the UDWT, 4 for the block partitioning, 4 for the discrete ridgelet transform.

Stack size	time (s)	memory (MB)
$512 \times 512 \times 32$	44.02	558
$512 \times 512 \times 16$	6.03	87
$256 \times 256 \times 128$	37.11	586
$256 \times 256 \times 64$	19.57	304
$256 \times 256 \times 32$	11.05	140
$256 \times 256 \times 16$	1.39	22

TAB. 4.1: Execution time results for the CVT3D forward transform and the corresponding memory requirement for storing the decomposition coefficients.

Because of the high redundancy of the CVT1 dictionary and the difficulty of interpretation of its coefficients a second generation of curvelets (CVT2) have been proposed in [57]. Two new curvelet implementations are presented : via an unequispaced FFT and thanks to wrapping in the Fourier domain. The wrapping version is the easiest to understand and to implement : it roughly consists in a coronization of the discrete Fourier domain based on concentric squares and shears as shown in Figure 4.3. Each part, coined *wedge* is then wrapped to a squared shape by exploiting some periodicity properties, hence allowing to analyze the wedge with the inverse FFT. These two implementations are much faster than the CVT1 (6 to 10 times a standard FFT) and the redundancy is also significantly reduced ($\sim 8J$ with J the number of scales). We show in Figure 4.4 a) the reconstruction of few curvelet coefficients which results in segments of different sizes and orientations in the image domain. The discrete FFT of this reconstruction (Fig. 4.4 b) illustrates the coronization of the discrete Fourier domain and also shows that the partitioning is smooth : wedges are overlapping, hence resulting in an redundant representation in the curvelet domain.

The natural extension of the CVT2 dictionary to the third dimension is presented in [185] and coined 3D discrete curvelet transform (CVT3D). The transform provides a sparse representation for 3D signals with surface-like singularities of codimension one but otherwise smooth. As for the 2D discrete transform the Cartesian grid is divided in wedges which are oriented in three dimensions. The discrete transform has a redundancy of about 5 which is moderately low.

The CVT3D is an appealing solutions for 3D curvilinear structures representation, however the implementations provided in the Curvelab package for Matlab[1] is not suited to 3D stacks of microscopy images. Indeed, it requires a large dyadic size along the z-axis but padding the image is not a feasible solution in our case for three reasons :

1. The number of slices rarely exceeds a dozen in tracking applications because of the fast acquisition demand and the coarser resolution in the microscope axis direction. On the other hand the xy-plane can be large (typically 512×512 pixels), so the 3D Fourier domain coronization can result in important issues for such anisotropic volumes. It would be required to partition the Fourier domain in a different way along each direction.
2. Image padding can significantly increase the computational load in the case of large size images.
3. The memory requirement for storing the images and their representation coefficients exceeds the capacity of standard computers because of the padding and the redundancy of the representations.

---

[1] Curvelab 2.1.2 by Emmanuel Candès, Laurent Demanet, Lexing Ying on http ://www.curvelet.org

Image size	time (s)	memory (MB)
128×128	0.06	0.40
256×256	0.18	1.53
512×512	0.76	6.00
1024×1024	3.86	24.00

TAB. 4.2: CVT2 forward transform execution time and memory requirements for a decomposition in four scales.

These assessments are supported by a number of experiences which are summarized in Table 4.1. We have processed 3D volumes of various sizes with the Matlab implementation of the CVT3D with the setup 3 (see Annex A).
The limitations of the CVT3D lead us to use 2D representation dictionaries slice-by-slice for representing a 3D volume. We discard here the use of the CVT1 because of its computational demand and we adopt instead the CVT2 dictionary. We summarize in Table 4.2 the memory and computational demand of using the CVT2 with different image size on the setup 3 (see Annex A). Due to the low memory requirements we are able to process much larger images than with a 3D transform.

4.3.2 Decomposition experiments

Plant cells processing

We have experimented the separation of background structures and KOR1 particles in plant cells images with the MCA algorithm. For the sake of visualization a crop of $256 \times 256 \times 11$ voxels was extracted from the original data. As shown in Figure 4.5 a) and b) these images contains a very structured background on top of KOR1 particles signal. We have applied the MCA algorithm slice by slice to the stack of images with the UDWT dictionary as D_1 and the CVT2 as D_2. The trade-off parameter λ was set to $1/5\sigma_n$ and 100 iterations were done with an exponentially decreasing threshold toward δ_f. As shown in Figures 4.5 c)-f) the separation of spotty and curvilinear structures is well achieved by the MCA procedure : the small compartments signal is mainly retrieved by the UDWT dictionary while elongated structures are reconstructed in the CVT2 image.

Prions and nanotubes separation

The source separation procedure has been used in a biological study of prions trafficking [78] presented in Chapter 8. The separation of tunneling nanotubes and prion complexes has been achieved in 3D fluorescence image sequences even if they overlap. The use of the MCA technique therefore make possible to track prions while moving inside the nanotubes. Without the use of the MCA technique the automatic analysis of the sequences would not have been possible because of the confusing signal originating from the nanotubes.

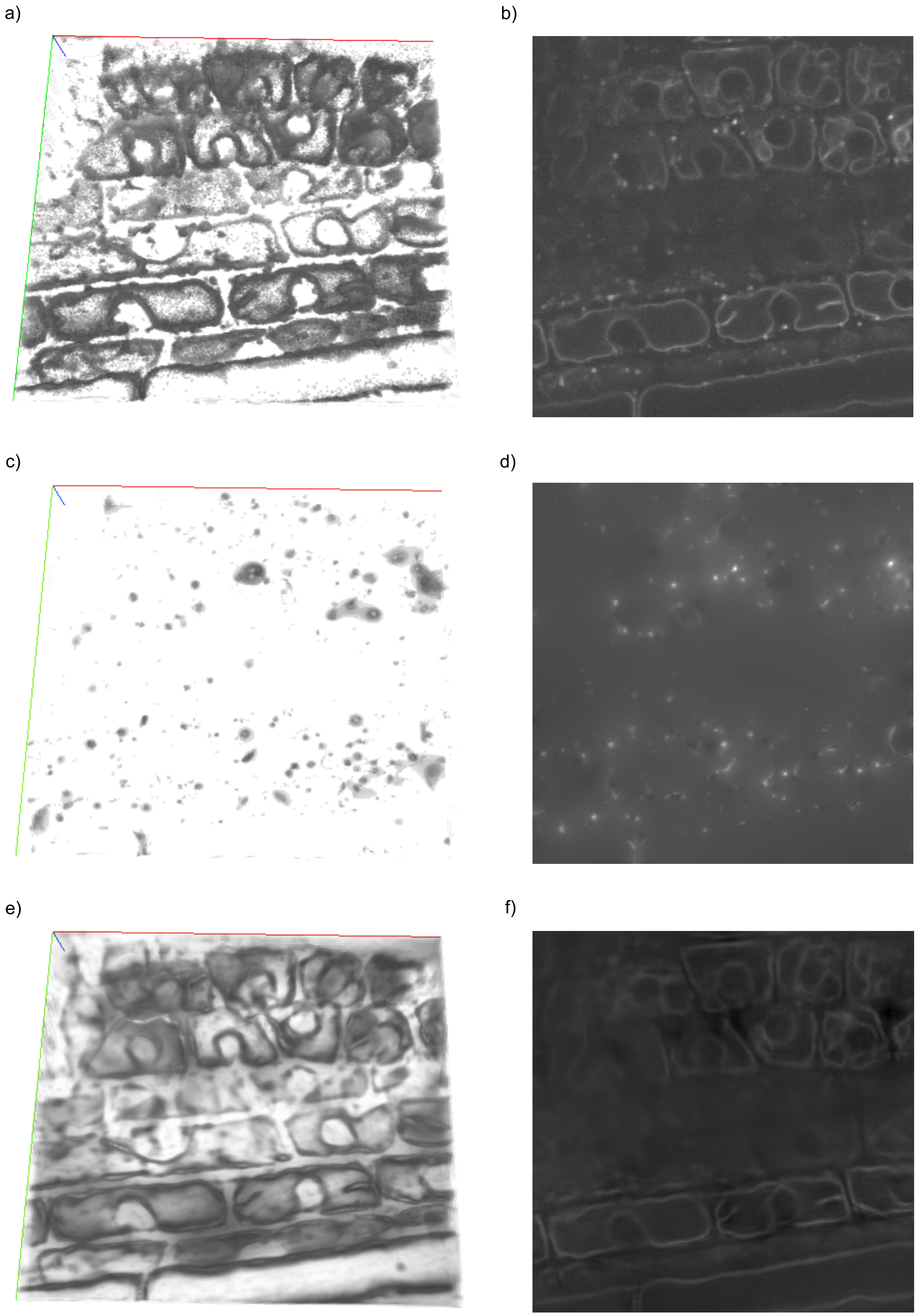


FIG. 4.5: MCA source separation on a 256×256 crop of plant cell images. a) Original stack of 11 images. b) One slice from the cropped stack. c) and d) Source s_1 recovered from the stack a) and the image b). e) and f) Source s_2 recovered from the stack a) and the image b).

4.4 Detecting and tracking fluorescent spots in complex backgrounds with MCA

4.4.1 Particle detection in a structured background

In order to detect particles in images containing a complex background we propose to first apply the MCA technique presented above to separate particle signal from background structures before applying a spot detection technique. We present next a detection procedure coined MCA-UDWT which relies on the analysis of the UDWT coefficients resulting from the MCA processing.

The MCA-UDWT detection scheme is summarized in Figure 4.6. We apply first the MCA procedure presented above on a sequence of fluorescence images and only wavelet coefficients $\hat{\alpha}_1$ are kept for the detection purpose. Wavelet scales corresponding to the size range of the targets are then extracted from $\hat{\alpha}_1$ and a spot-enhanced image is reconstructed by combining them. It is worth noting that we do not use any wavelet scale denoising procedure as in [187] since the denoising is embedded in the MCA procedure : wavelet coefficients have been selected by a hard thresholding which corresponds to a hypothesis testing procedure against noise. The spot-enhanced stack is reconstructed by putting the slices together, and particles 3D positions are estimated by detecting intensity maxima exceeding a predefined threshold.

We show in Figure 4.7 a zoom on a plant cell image and the sources retrieved by the MCA procedure. We note that some residues originating from background structures may remain in the target image. Hence, in order to avoid detecting these artifacts the first and second scales of the wavelet coefficients are extracted from $\hat{\alpha}_1$. Selecting scales allows us to specifically enhance particles, as shown in Figure 4.8. We also compare the original crop of image with the first and second scales of the B3-UDWT without any source separation procedure : Figure 4.8 shows that the UDWT filtering is not specific enough to enhance particles only and not the background. On the other hand the background is nearly cancelled when using the MCA scheme thanks to the competition which is imposed between the UDWT and the CVT2 for the background representation. We therefore expect a better performance of particle detection when using the MCA-UDWT than when using the UDWT alone.

In order to assess quantitative results, an expert identified 220 spots in the three first frames of the sequence corresponding to this crop. We compare the MCA-UDWT detection performance to those of a standard method consisting in thresholding the product of the first and second scales of UDWT coefficients without BSS, and then extracting detections in a similar way as described for MCA-UDWT. Figure 4.9 represents the number of recovered detections and the number of false detections for various thresholding strategies of the denoised image. We show that MCA-UDWT always recovers a higher number of targets than UDWT for a given number of false detections. When only few detections are allowed the MCA-UDWT is more sensitive than the UDWT, while when more false detections are allowed the MCA-UWT detects some membranes.

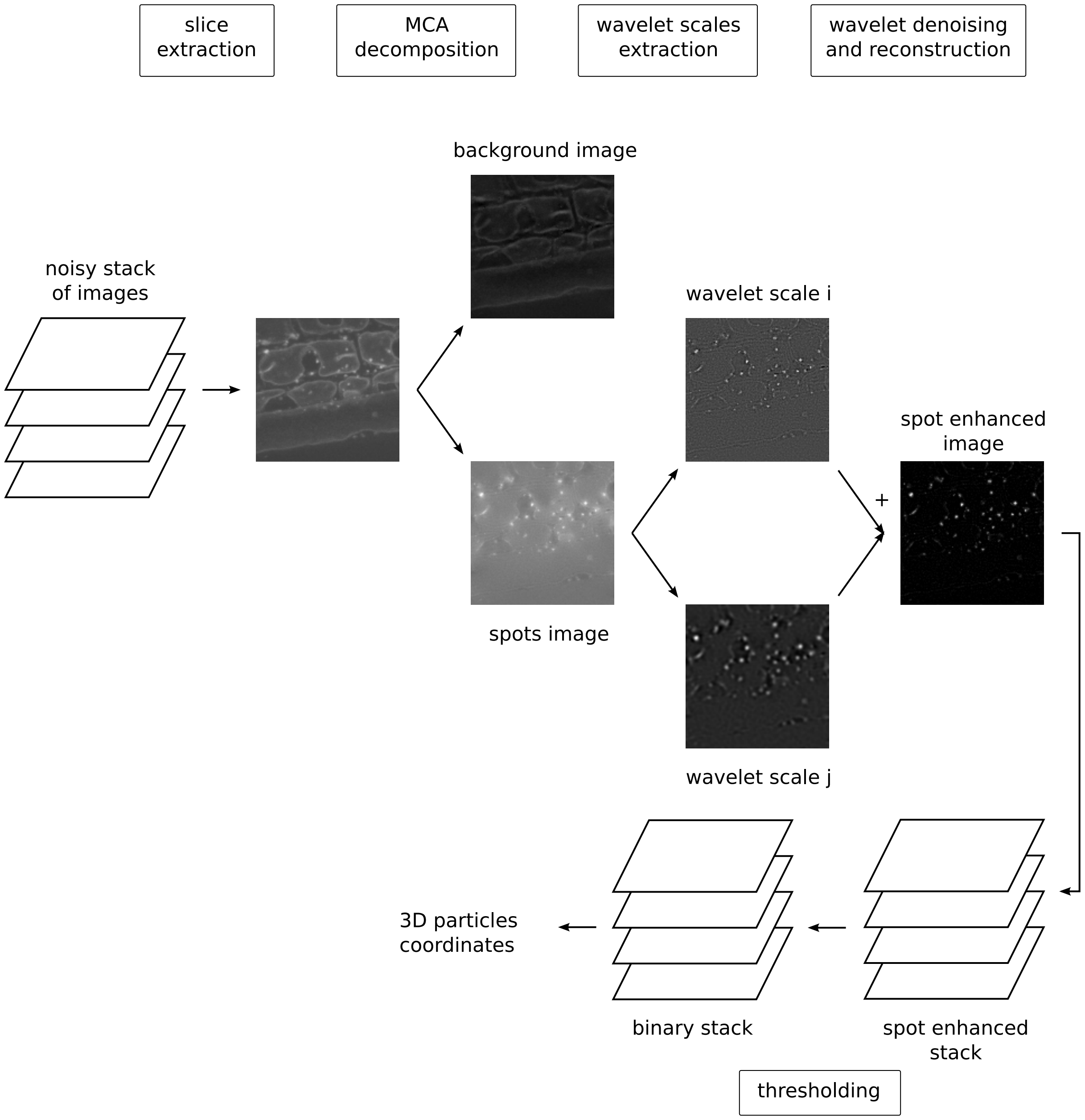


Fig. 4.6: MCA-UDWT spot detection scheme for a 3D stack of images containing a structured background. The main steps are : 1) extraction of a slice from the stack, 2) MCA decomposition of the slice, 3) extraction of the wavelet scales of interest, 4) reconstruction of a spot-enhanced image, 5) recombination of slices into a 3D stack, 6) thresholding of the stack, 7) extraction particles coordinates from the 3D binary mask.

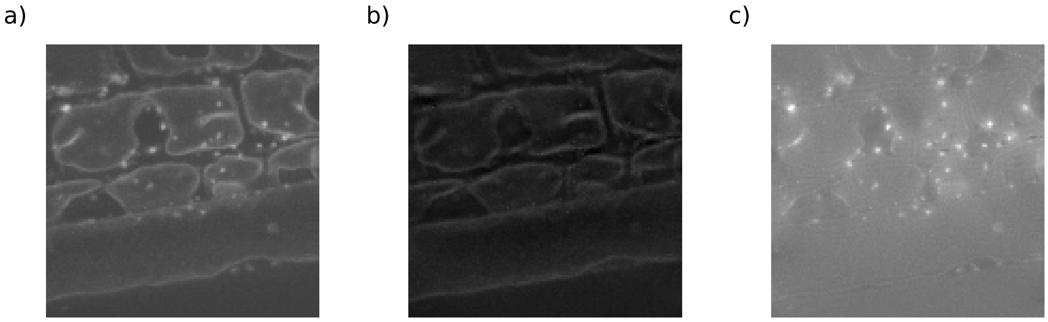


FIG. 4.7: Decomposition experiments on plant cell images. a) Crop of a 2D area of the original data. b) Background image extracted with the MCA. c) Particle image.

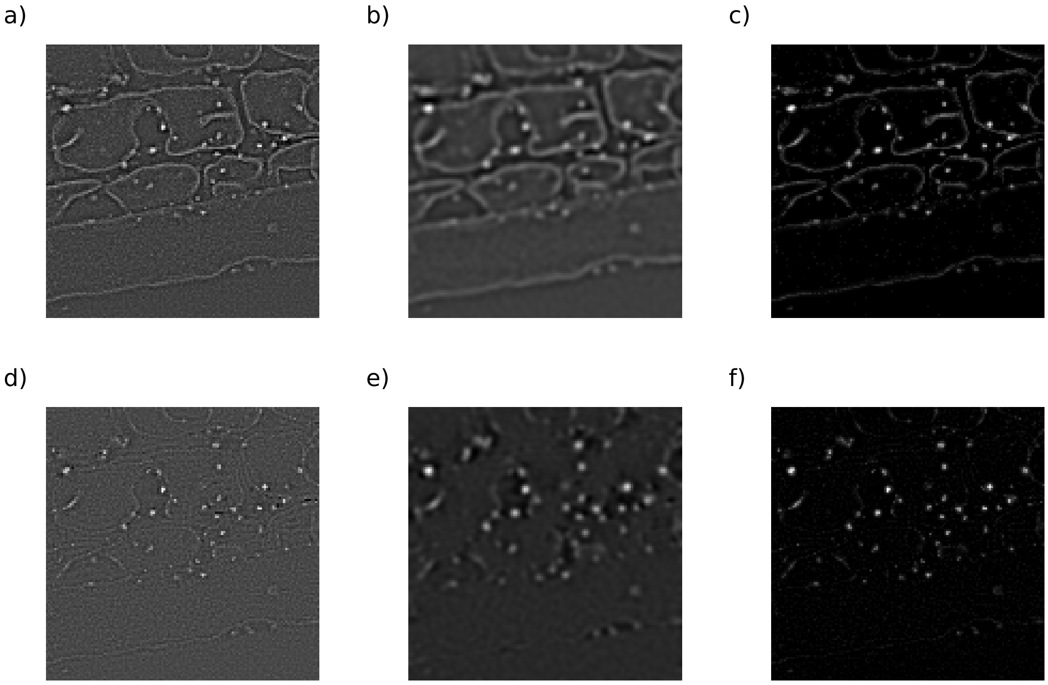


FIG. 4.8: Wavelet filtering on a small crop of plant cell images. a) and b) Second and third scales of the B3-UDWT applied to the plant cell image (Fig. 4.7 a)). c) Reconstruction using only scales 1 (a) and 2 (b). d) and e) First and second wavelet scales of the particle image extracted by the MCA (Fig. 4.7 c)). f) Image reconstruction using the first and second wavelet scales (images d and e respectively).

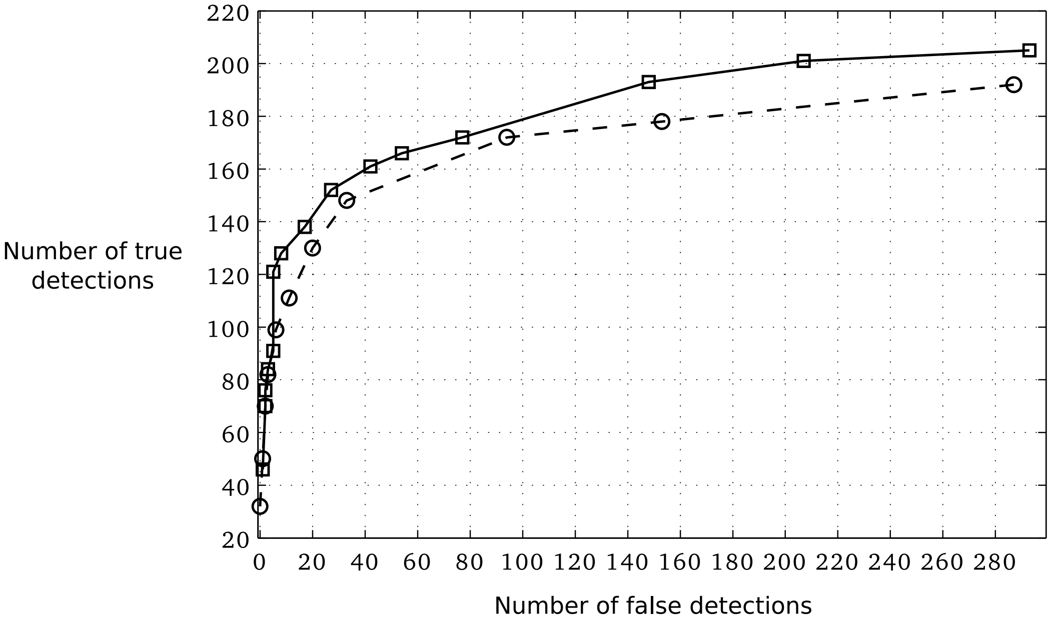


FIG. 4.9: Particle detection performance with and without source separation. The solid line is the detection performance for the MCA-UDWT technique with various thresholding strategies, while the solid line corresponds to the UDWT detection results.

4.4.2 Image augmented tracking

Once the detection step has been performed detections have to be linked together into tracks. Any point linking algorithm, such as the MHT presented in Chapter 7, can be used to process the detections. The simple additive Gaussian model (1.7) is however no longer valid after source separation, hence invalidating the use of tracking algorithms which relies on this model. As an example, the image augmented tracking algorithm presented in [39, 40] assumes a Gaussian shape of particles in the images. The distribution of the residual between images and the tracking model is compared to the noise statistics, hence making possible the separation of closely spaced targets. Such an approach is no longer valid after source separation due to the modified image model which no longer contains noise. We thus define next a procedure resulting in an image model suited to such tracking algorithms.

We estimate a background free image $\hat{I}$ by subtracting the background B to the measured image :

$$\hat{I}(x,y,z) = I(x,y,z) - B(x,y,w). \tag{4.13}$$

The background intensity $B(x,y,z)$ for the voxel (x,y,z) is in practice estimated thanks to the background source extracted by the MCA :

$$B = \mathbf{s}_2 = \mathbf{D}_2\alpha_2. \tag{4.14}$$

The background free image $\hat{I}$ can then be analyzed during the augmented tracking procedure since it still contains the particle signal T and noise term W.

In Figure 4.10 we show the results obtained by the image augmented tracking algorithm [40] when using the proposed modified image model (Eq. 4.13). The algorithm provides significant tracks despite the presence of the background in the original images. The proposed tracking procedure is indeed able to track particles interacting with membranes : this ability originates from the use of the BSS technique in the detection scheme. Indeed, without source separation detecting particles interacting with background structures is hardly feasible because their intensity values are hardly identifiable when superimposed to the background. We also note that no false tracks originate from some spurious detections produced by the background. Such false tracks cannot be avoided when detecting particles in the original image sequence because of the detection of background structures as objects of interest.

The modified image model we propose is also used to separate closely spaced targets that appear fused due to the resolution limit. The kinetic and image information are successfully mixed by the tracking algorithm to correct the fused detection. In Figure 4.11 the recovery of the aggregation of three particles is correctly processed despite the close proximity of membranes.

These results are opening new perspectives for biological studies : we are able to track particles in a very complex environment and also to quantify interactions with background structures since each separated image gives information about positions and dynamics of a source.

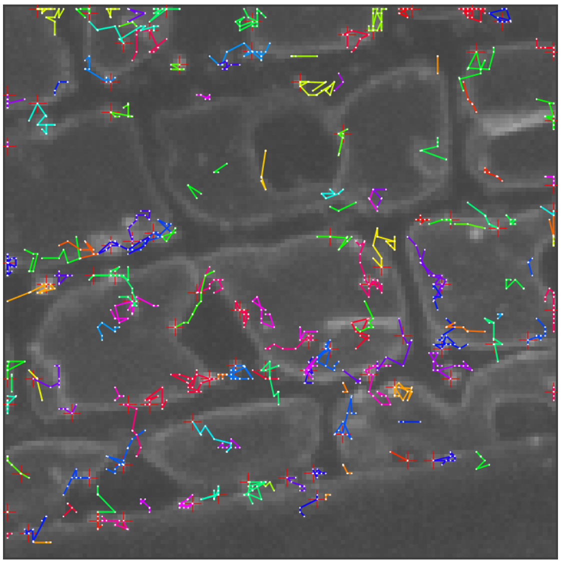


FIG. 4.10: Tracking numerous particles in plant cells by linking detections provided by the MCA-UDWT technique.

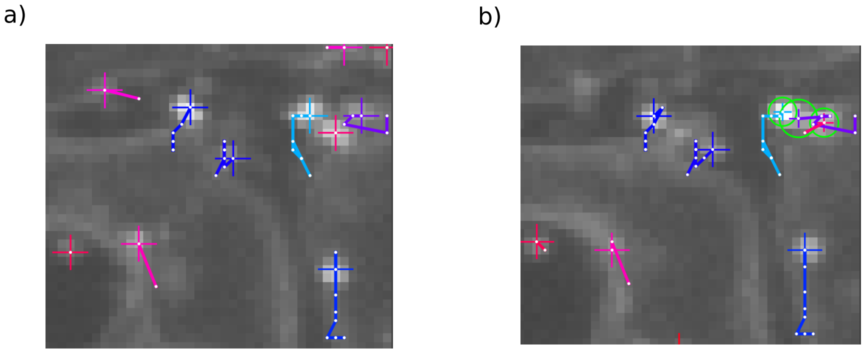


FIG. 4.11: Tracking particles in plant cells with an image augmented particle tracking procedure [40]. a) Detail of closely spaced particles at a given time k. c) Recovery of the aggregation of three particles at time $k+1$. Target profiles are represented by green circles.

4.5 Adapting the MCA to sequences of 3D fluorescence images

We have shown that using the a BSS technique is of great help for analyzing fluorescence images containing a structured background. The MCA algorithm combining the CVT2 and B3-UDWT represents an appealing solution to the BSS problem : a good signal separation can be achieved without much efforts. The separation procedure we have presented nevertheless still suffers from some limitations which degrade its performance. Adapting the MCA procedure to the characteristics of the 3D fluorescence images is a promising way to solve these issues. We present next some possible ways of improvement.

4.5.1 Positivity constraint for fluorescence image analysis

The intensity originating from fluorescent object is described by a Poisson random term of mean λ_p in the image model (Eq. 1.1) which indicates that intensity values due to their presence is positive :

$$\forall (x,y,z) \in I : \left\{ \begin{array}{l} T(x,y,z) \geq 0 \\ B(x,y,z) \geq 0 \end{array} \right. . \tag{4.15}$$

The positivity constraint is a strong prior on the two sources s_1 and s_2 which should be used when estimating them. This prior knowledge is however not used in the original MCA definition and negative intensity values are allowed. Allowing negative values in sources results in a significant degradation of sources estimation due to the oscillations produced by wavelet and curvelet coefficients reconstructions. In Figure 4.4 we give the reconstruction of few curvelet coefficients, and we show that near the high intensity segments some oscillations occur. In the MCA framework such oscillations are an important issue since other sources may produce ring artifacts compensating for them. Such ringing artifacts may complexify the particle localization and tracking procedures.

One solution to avoid ringing artifacts is to re-design image recontruction procedures from the coefficients $\{\hat{\alpha}_i\}_{i=1..2}$. Indeed, we use two overcomplete dictionaries D_1 and D_2 so there exists an infinite number of reconstruction schemes $\mathbf{D}'_1$ and $\mathbf{D}'_2$ as highlighted in [166]. An iterative undecimated reconstruction scheme has thus been proposed in [166] in order to avoid oscillating patterns around singularities and edges that are often created. The iterative procedure however requires an increased computation load which is not compatible with the MCA requirements. For the MCA purpose a simple and cheap idea to enforce positivity is introduced in [166] : negative values of a reconstructed source $\mathbf{D}_i\hat{\alpha}_i$ are replaced by some zeros in the alternate MCA scheme. Enforcing the positivity constraint in the BSS procedure reduces negative representation errors such as ringing artifacts and we therefore expect a better separation of sources and an easier localization of particles.

4.5.2 Sparsity priors

The MCA score function (Eq. 4.9) is intricately related to a Maximum-A-Posteriori (MAP) formulation [164] where the conditional probability of the representation α is written as :

$$p(\{\alpha_i\}_{i=1\ldots p}|Y) = \frac{p(\{\alpha_i\}_{i=1\ldots p})p(Y|\{\alpha_i\}_{1\ldots p})}{p(Y)}. \tag{4.16}$$

Following the source mixture model (Eq. 4.3) we compute the image conditional probability $p(Y|\{\alpha_i\}_{i=1\ldots p})$ as the probability of the residual between the observed image and the sources mixture. The residual should follow the same distribution as the noise N, hence under the assumption of white and Gaussian noise with zero mean and standard deviation σ_n we write :

$$p(Y|\{\alpha_i\}_{i=1...p}) = (2\pi\sigma_n^2)^{-\frac{n}{2}} \exp\bigg(- \frac{||Y - \sum_{i=1}^{p} \alpha_i D_i||_2^2}{2\sigma_n^2}\bigg). \tag{4.17}$$

Maximizing the log-likelihood therefore results in :

$$\begin{aligned} &\max_{\{\alpha_i\}_{i=1...p}} \ln(p(\{\alpha_i\}_{i=1...p}|Y)) \\ &= \max_{\{\alpha_i\}_{1...p}} \bigg[\ln(p(\{\alpha_i\}_{i=1...p})) - \frac{||Y - \sum_{i=1}^{p} \alpha_i D_i||_2^2}{2\sigma_n^2} \bigg]. \end{aligned} \tag{4.18}$$

Hence, by relating Equations 4.9 and 4.18 we deduce that minimizing the MCA score function is similar to maximizing the representation log-likelihood when :

$$\begin{aligned} &p(\{\alpha_i\}_{i=1...p}) = \beta^{-1} \exp(-\beta^{-1}||\alpha||_0), \\ &\text{and : } \lambda = \frac{\beta}{2\sigma_n^2}. \end{aligned} \tag{4.19}$$

The probabilistic distribution $p(\{\alpha_i\}_{i=1...p})$ defined in Equation 4.19 corresponds to an exponentially distributed sparsity prior of mean β. The choice of the exponential distribution is supported by the Maximum Entropy principle [95] : when the expected sparsity value β is the only prior information on the distribution of the coefficients α the exponential distribution of mean β is the maximum entropy distribution (the distribution which minimizes the amount of prior information built into the distribution). With the original definition of the sparsity pdf there is therefore no distinction between the coefficients corresponding to each source, and α is instead treated as a whole. However, in the standard MCA scheme (Algorithm 1) the representation in each dictionary $\hat{\alpha}_i$ is obtained by a hard thresholding scheme that corresponds to an exponential distribution of mean β for the prior pdf of the sparsity of the residual $\mathbf{R}_i$ in the dictionary D_i. In particle tracking applications specifying the prior distribution in each source may however be necessary. Indeed, we intend to separate particles, that are small, from background structures that generally form a much more complex signal. The redundancy of each dictionary is also different, so more coefficients should be allowed in a very redundant dictionary than in a few redundant one. In such a case the standard prior probabilistic distribution (Eq. 4.19) is not proper since it does not distinguish sources which can have very different sparsity values. We therefore propose to define as many prior distributions as the number of sources instead of defining only one prior :

$$p(\alpha) = \prod_{i=1..p} p_i(\alpha_i) \tag{4.20}$$

and :

$$p_i(\alpha_i) = \beta_i^{-1} \exp(-\beta_i^{-1}||\alpha_i||_0) \tag{4.21}$$

for $i = 1..p$. In Equation 4.21, the parameter β_i is the expected sparsity value for the coefficients α_i in the waveform dictionary D_i. Following the MAP formulation (Eq. 4.18) of the MCA inverse problem we rewrite the original score function :

$$\begin{aligned} &\min_{\{\alpha_i\}_{1...p}} \bigg[\sum_{i=1}^{p} \lambda_i^{-1} ||\alpha_i||_0 + ||Y - \sum_{i=1}^{p} \alpha_i D_i||_2^2 \bigg] \\ &\text{where : } \lambda_i = \frac{\beta_i}{2\sigma_n^2}, \text{ for } i = 1..p \end{aligned} \tag{4.22}$$

In order to maximize the proposed score function (Eq. 4.22) the MCA alternate projection/threshoding scheme can still be used by defining an independent thresholding strategy for each dictionary : the representation $\hat{\alpha}_i$ is obtained with a dictionary specific thresholding of the coefficients α_i. By doing so the prior distribution of the sparsity of the residual $\mathbf{R}_i$ in the dictionary D_i is implicitly assumed to be $p(\alpha_i)$ (Eq. 4.21), which is now correct.
In practice the prior sparsity means $\{\beta_i\}_{i=1..p}$ are not accurately known. We thus propose to estimate them by exploiting the continuity of images over time. Indeed, background structures and particles can move over time but their contribution to an image intensity is very similar between two successive images. We therefore assume that the sparsity of particles and the sparsity of the background are constant between two successive images. We propose to approximate the expected sparsity value β_i by $||\hat{\alpha}_i||_0$, the sparsity value measured in the previous image.

4.5.3 Other dictionaries for 2D and 3D curvilinear structures representation

A number of mutiscale and directional image decomposition schemes have emerged consecutively the development of CVT. The 2D contourelet transform (CT) [52, 59, 132] relies on a similar discrete Fourier domain coronization as the CVT2 but is based on a different implementation. The CT coefficients are obtained by applying a pyramidal decomposition (generally the Laplacian pyramid [32]) to the image, giving the multiscale property, and then analyzing the scales with a directional filter bank [8]. The CNT thus provides a high degree of flexibility both at the pyramidal decomposition step and in the choice of the filter bank. The redundancy obviously depends on the settings of these two steps : the Laplacian pyramid for instance induces a redundancy of 4/3 while both orthogonal and overcomplete dictionaries can be used in the filtering step. The approach has been extended to 3D with the surfacelet transform (SURFT) [115]. The implementation of the SURFT is based on successive filter banks which makes it fast even in 3D (~ 10 times the cost of a 3D FFT). The induced frequency tiling is still very similar to that of the CVT3D and the resulting redundancy is also low ($4 \sim 5$). The development of such methods is promising for the improvement of the source separation in 3D fluorescence images.

4.6 Conclusion

In this chapter we have proposed a novel approach to localize particles in fluorescence images containing a high intensity background. We base our detection technique on a prior separation of the images in two parts : the background part and the particle image. The separation is achieved thanks to the MCA algorithm which principle is to make different representation dictionaries compete for signals of different morphologies in the images. We have shown that using the MCA with adapted dictionaries provides a efficient way to separate the background from the particles. The resulting particle image is then analyzed to robustly estimate particles position. The results of the BSS technique can also be analyzed by sophisticated tracking procedures exploiting the image information to discriminate targets since we are able to erase the background signal in the original images. We have given a number of perspectives to adapt further the approach to fluorescence images of biological samples.

Chapter 5

Tracking multiple targets as a Maximum Likelihood problem

5.1 Maximum likelihood formalism

The tracking algorithms browsed through in Chapter 2 and 3 rely on either iterative track extension or track extraction from the whole spatio-temporal data, and in both cases the tracking is performed with respect to a score function which is local in some way. For instance the optimized score in grNN, IML and MTT methods is local to one track at a given time point since associations between a single track and a single detection are iteratively selected in a given frame. In contrast glNN and FPT are global in space, since the best association between the set of detections and the set of tracks is selected, but they are still local in time. Curve extraction methods [143, 25, 136] rely on a temporally global score function but which is spatially local since deciphering targets ambiguities is not computationally feasible in a global way in this framework. On the other hand, the gap capping proposed in [94] is global in both space and time, but relies on a preliminary NN tracking step and the global objective function does not find any simple and intuitive interpretation. The downside of using local score functions for target tracking is the lack of easy interpretation of the resulting trajectories : outputted tracks are optimal but in a undefined global sense. Hence, when changing the system setup such as imaging modalities, environment physical properties, or target types, score functions may change but in an unknown way when they are local to one frame or to one track. On the other hand, biological applications are so various that a high degree of flexibility and interpretability of methods is mandatory. We thus argue that the definition of a global score function for tracking with easy to interpret parameters is of prime importance.

5.1.1 Maximum likelihood as a global objective function

We propose here to consider the tracking task as the construction of the set of the most probable trajectories given the observed data. A similar objective function was introduced in the radar tracking community for the design of the MHT algorithm [138] and refined in [10, 9]. The set of trajectories built by the tracking procedure, denoted Θ^l, is composed of n tracks :

$$\Theta^l = \bigcup_{i=1..n} \theta_i. \tag{5.1}$$

A track $\theta_i \in \Theta^l$ is defined as a sequence of an arbitrary number of states at discrete time points t^k :

$$\theta_i = \{\mathbf{x}_i(t^{init_i}), \mathbf{x}_i(t^{init_i}+1), \ldots, \mathbf{x}_i(t^{end_i})\} \tag{5.2}$$

We denote here t^{init_i} the time of appearance in the surveillance volume of the i^{th} target and t^{end_i} the index of its last state frame. The time interval of its presence is denoted $\Delta_i = [t^{init_i}, t^{end_i}]$. By definition, t^{init_i} and t^{end_i} satisfy the following inequalities :

$$1 \leq t^{init_i} \leq t^{end_i} \leq l. \tag{5.3}$$

The state $\mathbf{x}_i(t)$ at any time time $t \in \Delta_i$ typically consists of the target spatial position, and can also contain information such as its velocity and features related to its appearance.
For tracking purposes we choose the likelihood of the tracks Θ^l as a score function to maximize. The likelihood is defined as follows :

$$\mathcal{L}(\Theta^l) = P\{\Theta^l | Z^l\}, \tag{5.4}$$

where Z^l is the observed data over the time sequence. The tracking task therefore consists in selecting the set of tracks $\Theta^{l,\star}$ which is the most likely in the set Ω^l of all the possible combinations of tracks :

$$\Theta^{l,\star} \triangleq \arg \max_{\Theta^l \in \Omega^l} \mathcal{L}(\Theta^l). \tag{5.5}$$

The tracking task is therefore defined as a standard *Maximimum Likelihood* (ML) problem.

5.1.2 Tracking by association of detections

Using intensities as measurements in the Bayesian framework has been proposed in [155] by using a standard Gaussian approximation with known size of particles' shape. The measurement model defined in the Bayesian framework is non linear in this case and non-linear estimation techniques such as SMC have to be used. This simple image model is however not valid in the most general case since particles can have different sizes and shapes with varying features. Moreover, in many biological applications the noise level is high compared to target intensity due to the low particle photo-emission and conditions such as cell auto-fluorescence and read-out noise as reported in Chapter 1. In this low SNR condition the image information is poor and not reliable enough, hence, techniques strongly relying on this data are of very limited benefit. Therefore, we adopt instead the standard two steps scheme : (1) particle detection, and then (2) positions linking, which is less sensitive to the imaging conditions and does not imply restrictive assumption on the particles' shape. In this case the measurements are just the detections which can be extracted either by a standard spot localization method (see Chapter 1) or with the spot detection method described in Chapter 4 when the background intensity is high.
We write the measurement at time k as the set of m_k detections in the k-th frame : $Z(k) = \{\mathbf{z}_i(k)\}_{i=1..m_k}$, where each measurement $\mathbf{z}_i(k)$ is composed of the detection's Cartesian coordinates : $\mathbf{z}_i(k) = [x_i(k), y_i(k), z_i(k)]^T$. Using a two steps procedure relying on detections yields a simplified linear measurement model which allows using fast linear estimation techniques in the state space model. When using detections as measurements, the tracking problem will be solved thanks to motion models of particles only. Including features such as appearance descriptors, the peak intensity or the size of spots is nevertheless easily feasible in the framework we describe,

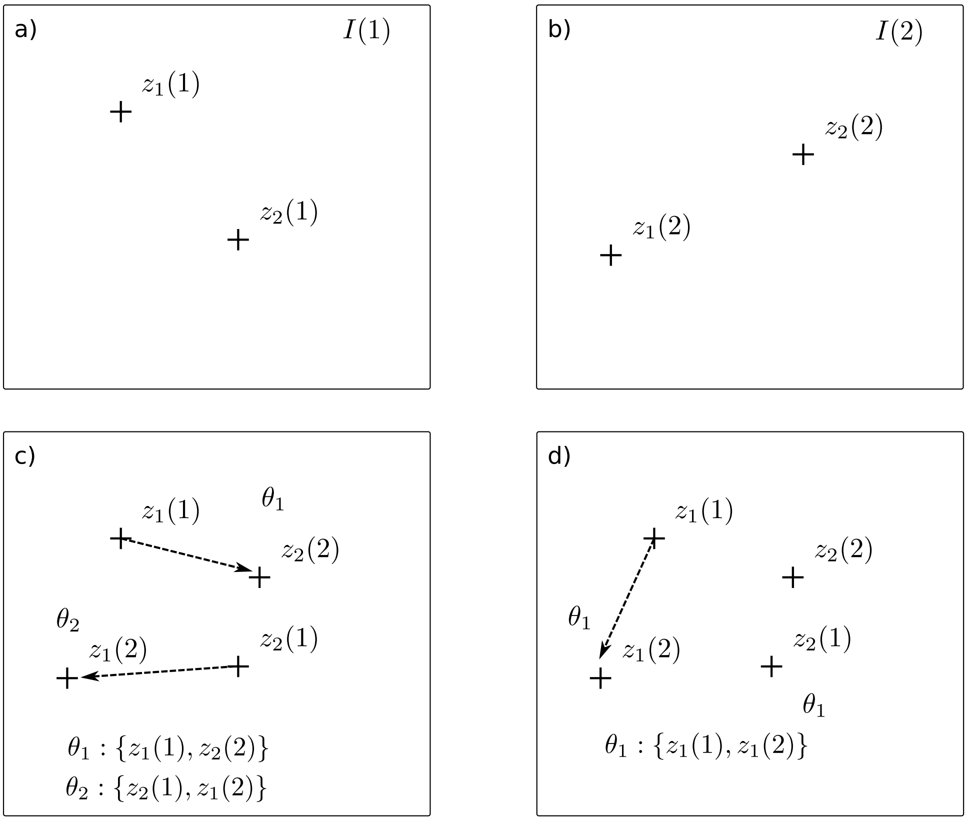


FIG. 5.1: Example of feasible associations for a simple system with two detections in each frame. Figure a) and b) shows detections extracted from two synthetic images $I(1)$ and $I(2)$. In Figures c) and d) two tracking solutions are presented. Two tracks are considered in the solution c), while only one track is considered in the case d) with two non associated detections. The exhaustive list of all the feasible solutions is given in Annex B : it shows that the large size of Ω^l prevents its enumeration for long image sequences containing several detections.

thanks to marginalization concepts : the idea is to split the pdfs in the Bayesian framework by assuming that targets motion and non-kinetic features behave independently.
As reported in Part I, the main limitation of the two steps tracking procedure is the corruption of the measurement procedure by false detections and missed detections in low SNR conditions. Hence, in contrast with state-of-the-art techniques, we will include the characteristics of the detection procedure in the probabilistic modeling of the tracking problem in order to solve issues inherent to such a corrupted measurement system.
The choice to account for kinetic information only when tracking is motivated by the need for an automatic procedure with a high degree of flexibility and thus applicable to all types of targets and imaging conditions. As shown in Chapter 8 we are able to track particles of very different types in various imaging modalities.

5.1.3 Solving the ML problem iteratively

In practice, enumerating all the feasible associations composing Ω^l is not feasible due to its large size : numerous associations result from the alternative associations between tracks and detections. The problem is also dramatically complexified by the lack of prior constraints on the possible tracks results. As an example, the number of targets is *a priori* unknown and new particles can appear over time. The number of solutions is moreover increased by the corrupted measurement process, that provides either significant or spurious detections, and which can miss some particles. We show for instance in Figure 5.1 a 2 frames example with four detections. In this fairly simple situation the number of tracking solutions is only 4 for one frame, but grows to 24 for Ω^1 (the detail of Ω^1 and Ω^2 is given in Annex B). Adding more frames and detections would result in an exponential explosion of the number of solutions which is emphasized by the possibility for particles to disappear in a number of frames.

In real cases, dealing with hundreds of frames with thousands of detections prevents the enumeration of the possible combinations of tracks. Dedicated schemes alleviating the construction of Ω^l have thus to be defined in order to solve efficiently the tracking problem. To do so, we first decompose the tracks likelihood $\mathcal{L}(\Theta)$ as :

$$\mathcal{L}(\Theta^l) = \frac{P\{\Theta^l, Z^l\}}{P\{Z(l)\}}, \tag{5.6}$$

by applying the Bayes rule. Since the probability $P\{Z(l)\}$ is constant across associations when solving Equation 5.5 the association problem can be re-written as :

$$\Theta^{l,\star} = \arg \max_{\Theta^l \in \Omega^l} P\{\Theta^l, Z^l\}. \tag{5.7}$$

We also make the usual assumption that particles move and are detected independently, so $P\{\Theta^l, Z^l\}$ can also be written as :

$$P\{\Theta^l, Z^l\} = P\{Z_0\} \prod_{j=1..n} p(\theta_j^{\Delta_j}, \mathbf{z}_{t_j}^{\Delta_j}), \tag{5.8}$$

where the couple $(\theta_j^{I_j}, \mathbf{z}_{t_j}^{I_j})$ corresponds to the set of positions of the target j and its associated measurements during the time interval I_j of its presence in the images. When a track is not assigned to a measurement from $Z(k)$ we associate it to a virtual measurement which is predicted by the dynamic filter of the target. In Equation 5.8 we note Z_0 the set of non associated measurements : these detections are considered to originate from clutter.

The independence of particle motion and detection results in the possibility to compute $\mathcal{L}(\Theta)$ as the product of independent probabilities (Eq. 5.8) ; however, finding $\Theta^\star$ cannot be achieved by maximizing independently the probability $p(\theta_j^{I_j}, \mathbf{z}_{t_j}^{I_j})$ for each track. Indeed, the association of detections from Z^l to the tracks imposes some strict constraints, such as the impossibility for two tracks to share a detection. The tracking problem has thus to be solved has a whole. To do so, we use again the Bayes rule to further decompose $P\{\Theta^l, Z^l\}$ as :

$$\mathcal{L}(\Theta^l) \propto P\{Z(l)|\Theta^l, Z^{l-1}\}P\{\Theta^l|Z^{l-1}\}, \tag{5.9}$$

and then as :

$$\mathcal{L}(\Theta^l) \propto P\{Z(l)|\Theta^l, Z^{l-1}\}P\{\Theta(l)|\Theta^{l-1}, Z^{l-1}\}P\{\Theta^{l-1}, Z^{l-1}\}, \tag{5.10}$$

where $\Theta(l)$ is the set of additional track states at time l that are used to extend tracks from Θ^{l-1} to Θ^l. Equation 5.10 shows that the joint tracks and measurements probability can be computed as an update of the joint probability up to time $l-1$. The update step consists of a product with two factors : $P\{Z(l)|\Theta^l, Z^{l-1}\}$ which accounts for the probability of the measurements at time

l when tracks are extended to the l^{th} frame, and the probability $P\{\Theta(l)|\Theta^{l-1}, Z^{l-1}\}$ modeling the probability of extending the tracks from frame $l-1$ to l (from Θ^{l-1} to Θ^{l}). By repeating the update process from the first frame to the last one, the tracks likelihood is computed as a product of probabilities over time :

$$\mathcal{L}(\Theta^l) \propto P\{Z(0)|\Theta(0)\}P\{\Theta(0)\} \prod_{k=1..l} P\{Z(k)|\Theta^k, Z^{k-1}\}P\{\Theta(k)|\Theta^{k-1}, Z^{k-1}\}. \tag{5.11}$$

In Equation 5.11 $P\{\Theta(0)\}$ is the probability of the tracks found in the first frame of the sequence and $P\{Z(0)|\Theta(0)\}$ is the probability of the set of measurements in the first frame when the initial tracks $\Theta(0)$ are given.

Equations 5.11 and 5.10 allow computing the tracks likelihood in an iterative way : the probability of initial tracks and detections are computed in the first frame, then the first joint probability $P\{\Theta(0), Z(0)\}$ is updated with the next observations and track extensions in the frame 1. The procedure can be iterated over time until the end of the sequence is reached. Relying on the iterative computation of the likelihood we propose a tracking algorithm in Chapter 7 which explores a subset $\tilde{\Omega}^l \subset \Omega^l$ containing $\Theta^{l,\star}$ with high probability. The key idea is to select only few tracking hypotheses at time k on the basis on the joint probability $P\{\Theta(k), Z(k)\}$, and then extend them with few association hypotheses in the next frame $k+1$ on the basis of $P\{\Theta(k+1), Z(k+1)\}$. When the partial hypotheses contained in $\tilde{\Omega}^k$ have a sufficiently high likelihood, the best hypothesis $\Theta^{l,\star}$ is indeed likely to be built from the extension of one of these hypotheses.

We detail in the next section the computation of track transition probability $P\{\Theta(k)|\Theta^{k-1}, Z^{k-1}\}$ which relies on statistical properties of particle motion, and we present in Section 5.5 corresponding measurement models allowing to compute the measurement conditional probability $P\{Z(k)|\Theta^k, Z^{k-1}\}$.

5.2 Association probability

We assume, as usually done, that the probability of the state of a target does not depend on previous measurements when the previous states are known. Hence, the tracks transition probability can be re-written as :

$$P\{\Theta(k)|\Theta^{k-1}, Z^{k-1}\} = P\{\Theta(k)|\Theta^{k-1}\}, \tag{5.12}$$

which only involves the previous target states. Some states in $\Theta(k)$ correspond to existing tracks that are extended from frame $k-1$ to frame k, while some correspond to new tracks which are observed for the first time in the frame k. The set $\Theta(k)$ is therefore split in two subsets $\Theta_0(k)$ and $\Theta_1(k)$ defined as follows :

$$\begin{aligned} &\Theta(k) = \Theta_0(k) \cup \Theta_1(k) \\ &s.t\ \forall \theta_j(k) \in \Theta(k) : \begin{cases} \theta_j(k) \in \Theta_1(k) \text{ if } \theta_j(k-1) \text{ exists in } \Theta(k-1), \\ \theta_j(k) \in \Theta_0(k) \text{ else,} \end{cases} \end{aligned}$$

and we denote here $J_1(k)$ the set of indexes in $\Theta(k)$ of tracks belonging to $\Theta_1(k)$, and $J_0(k)$ the set of indexes of new tracks.

We assume that the appearance of targets is independent of the presence of other targets and conversely, so the conditional probability is re-written as :

$$P\{\Theta(k)|\Theta^{k-1}\} = P\{\Theta_0(k)\}P\{\Theta_1(k)|\Theta^{k-1}\}, \tag{5.13}$$

which highlights the fact that the probability of the appearing targets can be computed separately from the transition of existing tracks.
Assuming that particles evolve in an independent manner, the conditional probability of the set $\Theta_1(k)$ is written as the product of each track transition :

$$P\{\Theta_1(k)|\Theta^{k-1}\} = \prod_{j \in J_1(k)} p(\mathbf{x}_j(k)|\mathbf{x}_j^{k-1}), \tag{5.14}$$

where $\mathbf{x}_j(k) = [x_j(k), y_j(k), z_j(k)]^T$ is the vector of the Cartesian coordinates of the j-th target at time k. We compute the transition probability for each target using the state space representation (Chapter 3) which offers a flexible framework to model the complexity of particle motion. The design of the state vectors and matrices is however a delicate task which depends on the state evolution to model and the known and unknown parameter values. Indeed, the state model have to reflect accurately particles motion in order to evaluate likelihood values which are consistent with the physical properties of particles. A dedicated model to particle motion is thus required in the state space framework.

5.3 Physical models of particle motion

Unlike vehicle, persons and cells, particles do not move on their own, but instead their motion is the unique consequence of local and global forces that drive them. Thus, we base the definition of kinetic models used for transition probabilities computation on physics of particles in fluids, which is a stochastic process long studied.

5.3.1 Brownian motion

Freely diffusing particles in a fluid have long been known to exhibit a Brownian motion. Indeed, as early as 1785 Ingenhousz observed the very erratic movement of coal particles in alcohol. The same experiment was repeated numerous times by other scientists such as Buffon and Brown, whose study of the motion of suspended particles of pollen in water gave its name to the Brownian motion. First descriptions of the Brownian motion characterized it as being very irregular and unpredictable with some properties linked to the medium temperature and viscosity, to the size of the particle but not to its nature. The formal model of the Brownian motion came later with the work of Einstein in 1905, who has shown that for a Brownian particle at position $\mathbf{x} = [0, 0, 0]^T$ at time 0 the pdf $P(\mathbf{x}, t)$ of finding it in position $\mathbf{x}$ at time t follows the diffusion equation :

$$\frac{\partial}{\partial t} P(\mathbf{x}, t) = D \Delta P(\mathbf{x}, t), \tag{5.15}$$

where Δ is the Laplacian operator and D is the diffusion constant. D is proportional to the media temperature T and inversely proportional to the particle's mass m according to the Einstein-Smoluchowski equation :

$$D = \frac{\kappa_B T}{m\gamma}, \tag{5.16}$$

where k_B is the Boltzmann constant and γ the absorption coefficient of the medium ($m\gamma$ is also called friction coefficient).
The probabilistic distribution of the particle position given by Equation 5.15 can be demonstrated

by the study of a well known process called the *Random walk* [182]. In the three dimensional Cartesian coordinates system, the diffusion equation admits as a solution the following pdf :

$$P(\mathbf{x}, t) = \frac{1}{(4\pi Dt)^{3/2}} \exp^{-\frac{||\mathbf{x}||_2^2}{4Dt}}, \tag{5.17}$$

which is a three-dimensional Gaussian density with a mean $\mu = [0, 0, 0]^T$ and a covariance matrix $\mathbf{\Sigma} = 2Dt\mathbf{I}_3$ (where $\mathbf{I}_3$ is the identity matrix of dimension 3). The zero mean indicates that the particle stays at the origin in average and the covariance matrix shows that position fluctuations grows in time as $t^{1/2}$. The position pdf (Eq. 5.17) can be extended in a straightforward way to a particle found at an arbitrary position $\mathbf{x}(k)$ at time k :

$$P(\mathbf{x}(k+\delta t)|\mathbf{x}(k)) = \frac{1}{(4\pi D\delta t)^{3/2}} \exp^{-\frac{||\mathbf{x}(k+\delta t)-\mathbf{x}(k)||_2^2}{4D\delta t}}, \tag{5.18}$$

due to shift invariance properties in space and in time. We denote next $\sigma_b^2 \triangleq 2D\delta t$ the variance of this normal position pdf which corresponds to a Brownian motion with a diffusion coefficient D.

5.3.2 Particle in a force field

Particles in biological environments are often not just freely diffusing as particles in water : instead a complex network of forces is acting on them on top of the thermal diffusion. If we suppose that a particle diffuses in a force field $\mathbf{G}$, then the Smoluchowski equation :

$$\frac{\partial}{\partial t} P(\mathbf{x}, t) = -\frac{1}{m\gamma} \nabla(\mathbf{G}(\mathbf{x}) P(\mathbf{x})) + D\Delta P(\mathbf{x}, t), \tag{5.19}$$

where ∇ is the divergence operator and γ a constant, characterizes the pdf of finding the particle at position $\mathbf{x}$ at time k in a similar way as the diffusion equation (5.15) for freely diffusing particles [139]. If the force field is constant : $\mathbf{G}(\mathbf{x}) = \mathbf{a}$ or harmonic : $\mathbf{G}(\mathbf{x}) = -a\mathbf{x}$, it can be shown that the solution to the Equation 5.19 is, in 3D, a Gaussian distribution :

$$P(\mathbf{x}, t) = \frac{1}{(2\pi|\Sigma(t)|)^{3/2}} \exp^{-\frac{1}{2}(\mathbf{x}-\mu(t))\Sigma(t)^{-1}(\mathbf{x}-\mu(t))^T}, \tag{5.20}$$

where the mean $\mu(t)$ and the covariance matrix $\mathbf{\Sigma}(t)$ are time varying and both depend on the force field $\mathbf{G}$. As for a diffusion the pdf can be easily modified for particles at an arbitrary position $\mathbf{x}(k)$ at an arbitrary time k. For a particle in a constant force field $\mathbf{G}(\mathbf{x}) = \mathbf{a}$ the two first moments of the pdf are :

$$\mu(k+\delta t) = \mathbf{x}(k) + \frac{\mathbf{a}}{m\gamma}\delta t, \tag{5.21}$$

$$\Sigma(k+\delta t) = 2D\delta t\mathbf{I}_3 = \sigma_b^2\mathbf{I}_3. \tag{5.22}$$

which shows that the average position is uniformly translated over time in the direction of the force field $\mathbf{G} = a$. Interestingly, the covariance matrix does not depend on the force field but is linked to the diffusion coefficient D only, as for the Brownian motion.

5.4 State transition models

The probabilistic description of the motion of particles in fluids provides physically based pdfs which are appealing for our particle tracking model. Indeed, in biology particles often diffuse in a very busy environment which could be modeled as a force field. The Smoluchovski equation (Eq. 5.19) is therefore well suited for describing the random diffusion of such particles. Characterizing the corresponding unknown force field is however a delicate task since biological environments, such as the cell, are generally very heterogeneous in space and changing in time. A very complex set of forces is therefore applied to diffusing particles, which is *a priori* unknown. We therefore propose to model the force field $\mathbf{G}$ as being random in space, time, and module.

We assume that particle movements are small compared to the force field fluctuations in space and that the force field changes slowly compared to the acquisition rate. We thus approximate the force field $\mathbf{G}$ acting on a particle on the path linking positions $\mathbf{x}(k)$ and $\mathbf{x}(k+1)$ by a constant field. In this case we use the Smoluchovski's description of the diffusion with a force field $\mathbf{G}(\mathbf{x}) = \mathbf{a}$ which is constant in a small neighborhood of $\mathbf{x}(k)$ during the time interval $(k, k+1)$. As shown in Section 5.3, the pdf of the position $\mathbf{x}(k+1)$ of the particle is a three-dimensional Gaussian distribution with mean $\mathbf{x}(k) + \frac{\mathbf{a}}{m\gamma}\delta t$ and a covariance $\sigma_b^2 \mathbf{I}_3$. In the following we denote $\mathbf{d}(k) = \frac{\mathbf{a}}{m\gamma}\delta t = [d_x(k), d_y(k), d_z(k)]^T$ the displacement vector which results from the effect of the force field $\mathbf{G}$ between frames k and $k+1$, leading to $\mathbf{x}(k) + \mathbf{d}(k)$ for the mean position in frame $k+1$.

In practice the force field $\mathbf{G}$ is unknown and an accurate estimate of its value is not available beforehand. Thus, for the sake of simplicity we model $\mathbf{G}$ as being constant during particle motion but we model the random changes between two frames in order to update our estimate $\mathbf{d}(k)$ in each frame k. We propose to include $\mathbf{d}(k)$ in the particle state vector as follows :

$$\mathbf{x}(k) = [x(k), y(k), z(k), d_x(k), d_y(k), d_z(k)]^T, \tag{5.23}$$

in order to estimate the displacement $\mathbf{d}(k)$ due to the force field action jointly with the position of the particle. By doing so the uncertainty on the force field value is accounted for when estimating a target position and the pdfs required for tracking. The variety of biological environments however requires the definition of dedicated state transition models including the force field estimation. We present next two such models.

5.4.1 White and Gaussian force field fluctuations

We suppose here that the force field $\mathbf{G}$ is locally constant and varies slowly in space and time. In the absence of information we assume a simple white and zero mean Gaussian model of force field fluctuations between two subsequent positions of a target which allows us to easily include the force field fluctuations in the state space representation : field fluctuations are part of the process noise. By doing so we are able to define a linear state transition model (Eq. 3.11) whereby the effect of the force field $\mathbf{G}$ and the thermal diffusion of the particle position are accounted for and the fluctuations of the force field modeled. Indeed, according to the Smoluchovski equation and the Gaussian model of force field fluctuations we can write :

$$\mathbf{F}_{\mathbf{Force}} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{Q}_{\mathbf{Force}} = \begin{bmatrix} \sigma_b^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_b^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_b^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & d_{xx}^2 & d_{xy}^2 & d_{xz}^2 \\ 0 & 0 & 0 & d_{yx}^2 & d_{yy}^2 & d_{yz}^2 \\ 0 & 0 & 0 & d_{zx}^2 & d_{zy}^2 & d_{zz}^2 \end{bmatrix}, \tag{5.24}$$

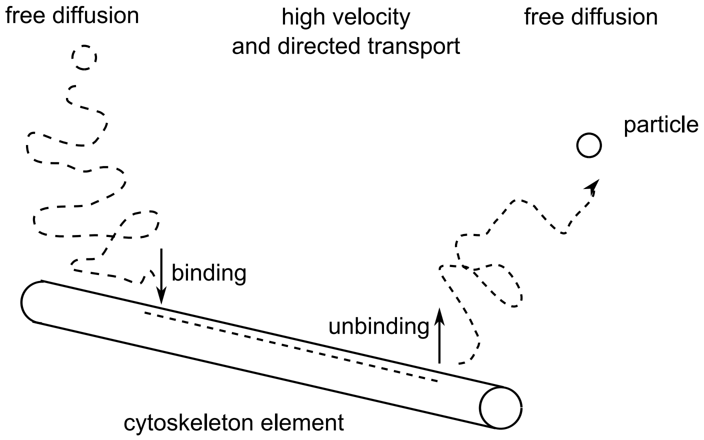


FIG. 5.2: The conveyor belt model principles. A particle freely diffuses in the media and attempts to bind to a cytoskeleton element with probability P_{on}. When the binding process succeeds the particle moves along the cytoskeleton element with a regular velocity. At each time step the particle attempts to unbind with probability P_{off}. Once unbound the particle starts to diffuse again.

for the state transition matrix and transition covariance matrix, respectively. The linear state transition model defined above allows using the KF estimation technique when a linear measurement model is chosen.

5.4.2 Cytoskeleton interactions modeling with multiple force fields

In cells, interactions of particles with cytoskeletal elements have been reported to give birth to a motion, coined the conveyor-belt motion [144, 146, 26], which alternates between free diffusion and directed transport by cytoskeletal structures. Cytoskeletal elements are assumed to form a spatially uniform conveyor belt moving with a fixed velocity. When a particle is bound to the conveyor belt, it moves at the velocity of the conveyor belt, and if it is not bound, it diffuses freely. At each time step, the particle binding or unbinding is a random event with given probability during the time interval δt : if the particle is free, it attempts to bind with probability P_{on} ; if it is not bound it attempts to unbind with probability P_{off}. The conveyor belt principle is illustrated in Figure 5.2.

The conveyor-belt motion corresponds to abrupt changes of the force field applied to a particle : the force field is nearly zero and random in direction when the particle is unbound, while it is large and highly directed when the particle is bound. The pdf of the force field is thus highly multimodal, in contrast with the unimodal Gaussian model assumed above. Moreover, the diffusion component is different in these two states : binding results in a tethering to the transporting structure which decreases the diffusion component. The white and Gaussian model of force field fluctuations is therefore not suited to the abrupt changes encountered with the conveyor-belt motion. Instead, we propose to use multiple models of force field and diffusion and to define a model of probabilistic switch between them to mimic the binding and unbinding events. We define a jump Markov state transition model with two models : free diffusion for the unbound state, and diffusion under the effect of a strong force field for the bound state.

For the free diffusion motion, the particle state is defined as the Cartesian coordinates : $\mathbf{x}(k) =$

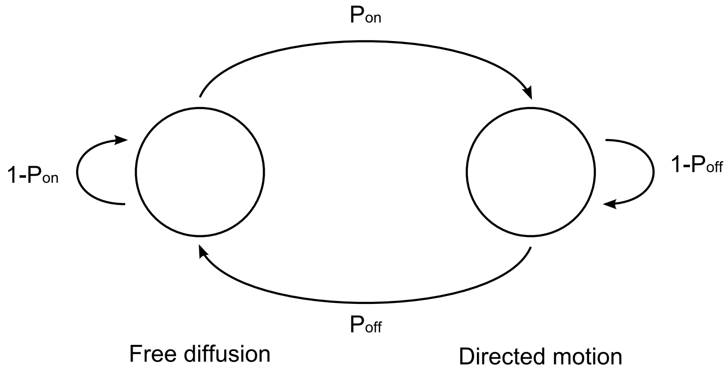


Fig. 5.3: First-order Markov chain for motion type switching between free diffusion and directed motion. The model switching states model mimics the conveyor belt model (Fig. 5.2) in which particles can bind and unbind to cytoskeleton structures with given probabilities. The bind and unbind probabilities : P_{on} and P_{off}, are used as transition probabilities in the HMM.

$[x(k), y(k), z(k)]^T$, whereby the state transition equation is written in a linear way with Gaussian process noise. The corresponding state transition matrices :

$$\mathbf{F}_{\mathbf{Brownian}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{Q}_{\mathbf{Brownian}} = \begin{bmatrix} \sigma_b^2 & 0 & 0 \\ 0 & \sigma_b^2 & 0 \\ 0 & 0 & \sigma_b^2 \end{bmatrix}, \tag{5.25}$$

are defined such that the conditional pdf of the position $\mathbf{x}(k+1)$ is a Gaussian distribution centered on $\mathbf{x}(k)$ as prescribed by the pdf (Eq. 5.18) characterizing the Brownian motion. For the bound state, we use the state and transition models defined for particle diffusing in a slowly varying force field (Eq. 5.24). The diffusion coefficient is set to a low value D_{bind} to account for the tethering to the cytoskeletal structure.

We model the motion type switching process with a first order Markov chain which is illustrated in Figure 5.3. We set the probability to switch from random walk to directed motion to P_{on} and the probability to switch from directed motion to Brownian motion to P_{off}, hence accounting for the true characteristics of the conveyor belt motion type. The probability of exhibiting the two kind of motions can be evaluated on line when observing a sequence of movements with standard Bayesian computation. An exact computation however requires storing every possible combination of states over the sequence, which is not practical. Indeed, the probability of a given measurement at time k does not only depend on the actual state motion, but on the whole motion states history which serves to estimate the actual target state $\mathbf{x}(k)$ and the pdf of its measurement. In practice, enumerating combinations of motion types is not feasible for long tracks since the number of state combinations is 2^l for a track of length l. We therefore simplify the model by adopting an IMM filter where combinations of motion types are iteratively merged, resulting in only one filter for each motion type with modified pdfs to account for motion types history (see Chapter 3 for a detailed discussion on the IMM estimation technique).

5.4.3 Appearance of new targets

We model the appearance of new targets in the stack as a stochastic process since these events depend on phenomena which are not accurately known in the general case. In the radar community the number of new targets is generally assumed to follow a Poisson distribution with a fixed mean, while the spatial distribution of theses events is supposed uniform in the whole

surveillance volume. We propose instead a model which gives us the capability to define more precisely the properties of the target appearance process and the influence of various physical properties of the environment.
We suppose that in every voxel $\mathbf{x} = \{x, y, z\}$ of the stack a target can appear with a probability $P_{NT}(\mathbf{x}, k)$ which is both time and location dependent in the most general case. Indeed, the biological environment may favor the entry of new targets in specific areas of the surveillance volume because of its topology. The rate of target appearance can also vary over time due to some biological events or some physical changes of the environment. The probability of a set of new tracks in frame k is therefore computed as :

$$P\{\Theta_0(k)\} = \prod_{j \in J_0(k)} P_{NT}(\mathbf{x}_j(k), k), \tag{5.26}$$

in the most general case. Without any specific information on the probability map of target appearance $P_{NT}(\mathbf{x}, k)$ we assume it constant in space and over time : $P_{NT}(\mathbf{x}, k) = P_{NT}$ for every position $\mathbf{x}$ in the stack in every frame. In this case the probability of the set of new tracks is simply computed as :

$$P\{\Theta_0(k)\} = P_{NT}^{|\Theta_0(k)|}. \tag{5.27}$$

The model of target appearance is also applicable to tracks in the first frame : the number of objects in the first frame is also a random variable, and since no past information is available we consider that every track in $\Theta(1)$ corresponds to an appearing target : $\Theta_0(1) = \Theta(1)$. In the first frame the probability for each voxel to generate a new target is set to a higher value than in other frames : $P_{NT}(\mathbf{x}, 1) > P_{NT}(\mathbf{x}, k)$ for $k = 2..l$.

5.5 Measurement model

5.5.1 Linear estimation

The measurement process consists in detecting the potential positions of particles in the current frame either with a standard spot localization method (Chap. 1) or the spot detection technique presented in Chapter 4. Each active track $\theta_j(k)$ in frame k can be associated to one measurement $\mathbf{z}_{t_j}(k)$ by the tracking procedure. The measurement process thus provides noisy Cartesian coordinates for active targets, which yields an independent linear measurement equation (Eq. 3.10) in the state space since state models presented above include the target coordinates in the same Cartesian system. The corresponding measurement matrix for each target is specific to each motion model :

$$\mathbf{H}_{\mathbf{Force}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \tag{5.28}$$

for particles diffusing in a force field with slow white and Gaussian fluctuations, and :

$$\mathbf{H}_{\mathbf{Brownian}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{5.29}$$

for particles exhibiting a pure Brownian motion.
The measurement noise corresponds to the random inaccuracy of the detection procedure, which is usually assumed to be Gaussian. We thus write the measurement noise covariance matrix as :

$$\mathbf{R}(k) = \begin{bmatrix} \sigma^2_{m,xy}(k) & 0 & 0 \\ 0 & \sigma^2_{m,xy}(k) & 0 \\ 0 & 0 & \sigma^2_{m,z}(k) \end{bmatrix}, \tag{5.30}$$

when assuming that the errors of localization along each of the three directions are independent. The localization error is usually larger along the z-axis than in the $xy-$plane ($\sigma^2_{m,z} > \sigma^2_{m,xy}$) because of the coarser resolution in depth. The accuracy of the detection procedure is intricately linked to the detection procedure which is used and is also highly SNR dependent : the amplitude of the errors grows as the SNR decreases since the particle signal is increasingly obscured by the random noise. As reported in Chapter 1 the SNR is in general time dependent in fluorescence image sequences because of the photobleaching effect. Tuning the localization variances $\sigma^2_{m,xy}(k)$ and $\sigma^2_{m,z}$ to the current frame k could thus be necessary. Doing so is easily achievable by simulating particles in a synthetic image reproducing the desired SNR conditions. In practice the localization accuracy can be lower than 1/10 of a pixel in the $xy-$plane in moderately low SNR conditions (PSNR$\sim$ 10) with standard localization techniques presented in Chapter 1.
The measurement generation process for one particle can be considered independent of other targets presence and detections. Therefore, as the state transition model is also independent for each particle, the state estimation can be achieved independently for each particle. For each type of motion the state space equations are linear and corrupted by a white Gaussian noise, so we use the KF which is computationally cheap and optimal in this case, as seen in Chapter 3. For particles transported by cytoskeletal structures we use an IMM filter embedding two KFs : one for Brownian diffusion and one for motion under a force field effect.

5.5.2 False detections probabilities

The set of non associated measurements $Z_0(k)$ in each frame k is classically assumed to correspond to detections produced by artifacts since they do not originate from a target. The number of such detections and their position in images are therefore random. In the radar tracking literature [20] the number of such false detections is generally assumed to follow a Poisson distribution, while detections are assumed to be uniformly distributed in the whole volume. We prefer instead to define a specific false detections model which is tuned to the characteristics of the localization of fluorescent particles in microscopy images.
As shown in Chapter 1, particle detection methods often rely on a spot enhancing step, followed by a pixel-wise intensity thresholding to obtain a binary image of spots. When processing real images each pixel of the filtered image can randomly exceed the detection threshold. We note P_{FD} the probability of such event. In practice the probability of emitting false detections can be time and position dependent. Indeed, biological environments are generally very heterogeneous, hence the background is uneven and some parts of the image which are not of interest may be highly fluorescent. Estimating the map of probability of generating false detections is required in such a case. It can be done experimentally by applying the detection algorithm to real or synthetic data. The probability of the set of detection is therefore written as :

$$P\{Z_0(k)\} = \prod_{\mathbf{z}_j(k) \in Z_0(k)} P_{FD}(\mathbf{z}_j(k), k), \tag{5.31}$$

in the more general case. When the background is supposed approximately uniform in both space and time the detection probability is assumed constant and the probability of $Z_0(k)$ is simply written as :

$$P\{Z_0(k)\} = P_{FD}^{|Z_0(k)|}. \tag{5.32}$$

Estimating the parameter P_{FD} is in this case pretty straightforward : it only requires generating empty synthetic images containing a noise reproducing the real SNR conditions, applying then the detection algorithm, and counting the number of measurements. The probability of generating false detections is estimated by the ratio of the mean number of detections λ_{FD} to the volume of the image V : $P_{FD} = \lambda_{FD}/V$. It is worth noting that some detection algorithms rely on a pixel-wise statistical hypothesis test to discriminate noise from targets. In such a case the probability of false detection P_{FD} corresponds to the error of the first kind of the hypothesis test.

5.6 Estimating model parameters

The state space representation presented above for particle tracking involves only a few parameters but which need to be set. Estimating the measurement covariance matrix is straightforward with detection experiments on synthetic images, however, motion models parameters can be more difficult to evaluate. More specifically, the process noise covariance matrix $\mathbf{Q}$ plays a key role in the computation of association pdfs but its value is not known *a priori*. We propose two approaches to evaluate $\mathbf{Q}$: (1) on line estimation of $\mathbf{Q}$ for each target from the observed movements, (2) automatic estimation of the diffusion coefficient in a pre-processing step. We describe these two solutions next.

5.6.1 On line diffusion estimation for individual particles

The main advantage of estimating the matrix $\mathbf{Q}$ while tracking is the possibility to discriminate particles which diffuse differently. Differences in the diffusion process may result from size and mass variations and imply a specific diffusion coefficient (Eq. 5.16) for each particle. We adopt here an approach similar as the procedure described in [65] : the observed movements are used to refine a likely covariance matrix $\mathbf{Q}$ to each frame and each particle. In practice the covariance matrix at time $k+1$ is estimated as :

$$\mathbf{Q}(k+1) = (1-\alpha)\mathbf{Q}(k) + \alpha[\mathbf{x}(k|k) - \mathbf{x}(k|k-1)][\mathbf{x}(k|k), -\mathbf{x}(k|k-1)]^T, \tag{5.33}$$

which corresponds to an update of the previous covariance matrix with respect to the difference between the predicted state and the estimated state for the current frame. The difference $[\mathbf{x}(k|k) - \mathbf{x}(k|k-1)]$ is called the *innovation* of the process and is often used to quantify the goodness of fit of the state space model : a suited states space model should indeed yield normally distributed innovations. The parameter $\alpha \in [0,1]$ in Equation 5.33 is a memory parameter, the role of which is to fix a trade-off between the robustness to movements with low probability and the rapidity of the adaptation to sudden changes of diffusion characteristics.

5.6.2 Practical evaluation of the on line estimation of the diffusion

In order to test the accuracy of the proposed procedure for the estimation of the diffusion process we have conducted two experiments : (1) estimation of the diffusion process for a simulated particle, (2) adaptation of the estimation scheme to an abrupt change of diffusivity.

Estimation of σ_b^2

We have first simulated the acquisition of 100 microscopy images showing a particle which diffuses in the xy-plane. We have set the covariance of the Gaussian process simulating the Brownian motion to : $\sigma_b^2 = 16\ pixels^2 \cdot frame^{-1}$. The process noise covariance matrix $\mathbf{Q}$ was initialized as : $\mathbf{Q} = 25 \cdot \mathbf{I}_3$. The unappropriated $\mathbf{Q}$ has been chosen here to mimic the lack of prior accurate description of the diffusion process. We check with this realistic scenario the capability of the estimation procedure (Eq. 5.33) to converge to the true value σ_b^2. We show in Figure 5.4 the estimated value of σ_b^2 resulting form the on line estimation of $\mathbf{Q}$ for three different α values : 0.05, 0.15, 0.5. For the three settings we verify that the empirically estimated value σ_b^2 fluctuates around the true value : σ_b^2. We however note some important differences : the extent of variations around the true σ_b^2 value is intricately linked to the α tuning. When α is low the newly observed innovation has a low weight in the estimation of $\mathbf{Q}$, hence the procedure slowly adapts to the true value σ_b^2, but the estimation process is robust to the jumps of the particle which have a low probability. A low α value indeed corresponds to averaging innovations from numerous previous frames. We observe in practice that approximately 50 frames are required for the algorithm to converge to σ_b^2 when $\alpha = 0.05$. Increasing α results in a higher weights for the current innovation, hence leading to a faster convergence to the true solution and a shorter impact of the initial $\mathbf{Q}$ setting : analyzing only two frames is required to reach the value $\sigma_b^2 = 16.81$ when $\alpha = 0.5$, while the value 16.65 is reached after five frames for $\alpha = 0.15$. Using a high α setting however yields a poor robustness of the estimation procedure against low probable movements : each very high or very low value of the innovation leads to an abrupt change of the estimation of σ_b^2, which may result in inconsistent probability calculations. Using a moderately low α value is therefore required to gain robustness against such events.

Adaptation to sudden changes of the characteristics of the diffusion process

The ability of the estimation procedure to adapt to diffusivity variations has also been investigated by simulating a sudden change of the σ_b^2 parameter for a diffusing particle. After 50 frames for which σ_b^2 is set to 4, we abruptly change it to the value 36. The covariance matrix process noise covariance matrix is initialized as : $\mathbf{Q} = 25 \cdot \mathbf{I}_3$. The estimation results are given in Figure 5.5 for $\alpha = 0.05$, 0.15 and 0.50. We show that the setting $\alpha = 0.05$ is not flexible enough to rapidly compensate for the error of initialization and to model the change of diffusivity : tens of frames are required to reach a satisfactory estimation in each case. By using a larger value for α the model self adapts rapidly to the new conditions. We nevertheless still remark that the larger the value of α, the less robust is the estimation procedure against events with a low probability.

The two scenarios analyzed in Figures 5.4 and 5.5 lead us to chose a moderately low value for the trade-off α between the robustness and the adaptivity of the estimation procedure. The setting $\alpha = 0.15$ is an appropriate choice which we commonly use.

5.6.3 Estimating the diffusion beforehand

Estimating motion characteristics on line allows one to account for the motion specificity of each particle and to adapt to fluctuations when tracking. However, when particles are very similar and no sudden physical changes of the media happen, diffusion characteristics are expected to be very similar across particles and over time. In this case the on line diffusion estimation procedure is of limited benefits and the solution may lack robustness against low probability events, as shown in Figures 5.4 and 5.5. Moreover, few frames are still required for the estimated

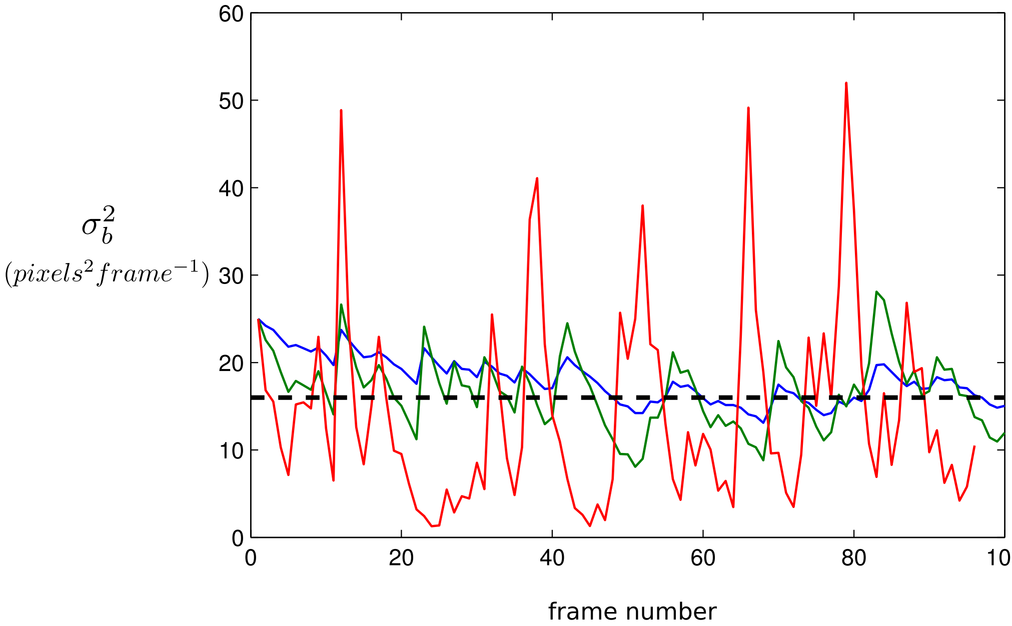


FIG. 5.4: Diffusion parameter σ_b^2 estimation. Three α strategies are compared : 0.5 (red), 0.15 (red) and 0.05 (blue). The initial σ_b^2 value is 25 and the true value is 16. A dashed line is plotted to represent this goal.

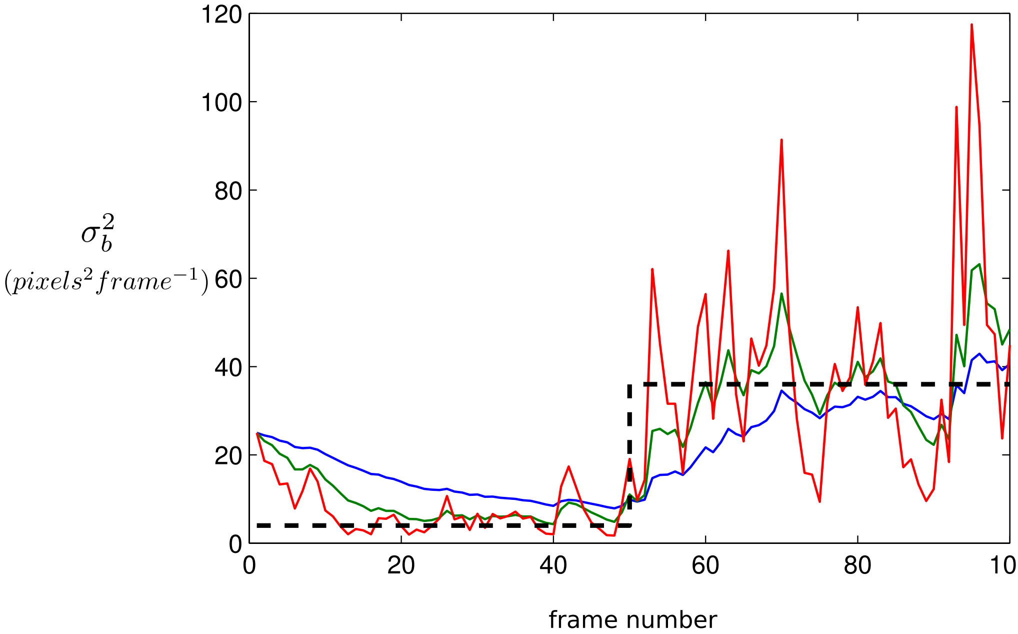


FIG. 5.5: Parameter estimation for abrupt diffusion changes. The diffusion parameter σ_b^2 exhibits a sudden change at frame 50 : it grows from 4 to 36 (dashed line). Three α strategies to estimate σ_b^2 are compared : 0.5 (red), 0.15 (red) and 0.05 (blue).

σ_b^2 to converge to the true motion characteristics.

We propose here a novel solution to these issues which consists in evaluating the diffusion characteristics beforehand, without tracking. When particles have little diffusion differences we still propose to use the procedure to provide a rough estimate of the diffusion which is then refined by the on line estimation procedure. The long period of adaptation to the true diffusion value which occurs when using a low α value (Fig. 5.4 and 5.5) can thus be alleviated by initializing $\mathbf{Q}$ with the provided approximation of σ_b^2.

Models of distance between detections

The distances travelled by particles between subsequent frames can be used to estimate the parameter σ_b^2 characterizing the diffusion process since the pdf of the position is given by a normal distribution involving σ_b^2 (Eq. 5.18). We note that in the majority of cases detections originating from the same target are the closest ones in two subsequent frames. The smallest distance between one given detection in the frame k and the set of detections in the frame $k+1$ can therefore give a hint about the diffusion characteristics. Inferring σ_b^2 is however not straightforward since particles are not actually tracked over time : we do not account for missed and false detections, nor do we model target appearance and disappearance. Hence a significant number of distances between detections do not correspond to a Brownian jump between the two subsequent frames. The key idea of the approach we propose is to analyze the histogram of distances between closest detections from subsequent frames to discriminate Brownian jumps and distances computed from false detections and wrong associations of detections.
The squared distance between two detections $\mathbf{z}_i(k)$ and $\mathbf{z}_j(k+1)$ in subsequent frames is computed in 3D as :

$$\begin{aligned} d_{ij}^2(k) &= [\mathbf{z}_j(k+1) - \mathbf{z}_i(k)][\mathbf{z}_j(k+1) - \mathbf{z}_i(k)]^T \\ &= (x_j(k+1) - x_i(k))^2 + (y_j(k+1) - y_i(k))^2 + (z_j(k+1) - z_i(k))^2. \end{aligned} \tag{5.34}$$

If the detections $\mathbf{z}_i(k)$ and $\mathbf{z}_j(k+1)$ originate from the same Brownian target the squared distance $d_{ij}^2(k)$ is computed as the sum of three squared terms (Eq. 5.34) each following an independent normal distribution due to the Brownian motion and the Gaussian measurement error. In the xy-plane the variance of the distance is $\sigma_b^2 + \sigma_{m,xy}^2$, and it is $\sigma_b^2 + \sigma_{m,z}^2$ along the z-axis. Hence the normalized distance $\hat{d}_{ij}^2(k)$ follows a chi-square distribution χ_d^2 whose degree of freedom d is the number of dimensions of a stack. For instance in 3D :

$$\begin{aligned} \hat{d}_{ij}^2(k) &= \frac{(x_j(k+1) - x_i(k))^2}{\sigma_b^2 + \sigma_{m,xy}^2} + \frac{(y_j(k+1) - y_i(k))^2}{\sigma_b^2 + \sigma_{m,xy}^2} + \frac{(z_j(k+1) - z_i(k))^2}{\sigma_b^2 + \sigma_{m,z}^2} \\ &\sim \chi_3^2. \end{aligned} \tag{5.35}$$

The pdf of $\hat{d}_{ij}^2(k)$ allows us to establish a direct link between the histogram of detection distances and the parameter σ_b^2.
In practice the detections set is corrupted by some false detections, some detections are missed, and appearance/disappearance events occur. Hence, some of the smallest distances between detections of subsequent frames do not correspond to jumps of Brownian particles. We propose two approaches for the estimation of the pdf of theses distances : (1) a theoretical estimation based on the uniform spatial distribution of detections, or (2) the empirical estimation of the pdf by the analysis of observed distances between detections in the same frame.

Probabilistic distribution of the distance between detections originating from different targets

Assuming that false detections and target appearance events are uniformly distributed in the surveillance volume V, we can derive a theoretical distribution for the smallest distances between detections that do not originate from the same target. We demonstrate in Annex B that the pdf of the minimal Euclidian distance d between a particle detection and p other uniformly distributed detections is :

$$f(d) = p\frac{4\pi d^2}{V}\left(1 - \frac{4\pi d^3}{3V}\right)^{p-1}, \tag{5.36}$$

where V is the volume of the stack. In the 2D case the minimal distance pdf is given by :

$$f(d) = p\frac{2\pi d}{V}\left(1 - \frac{\pi d^2}{V}\right)^{p-1}. \tag{5.37}$$

The assumption of an uniform spatial distribution of detections may not be fulfilled in some cases where the topology of the environment plays an important role. We thus propose that in such a case, an empirical pdf is evaluated for the minimal distances between detections that do not originate from the same particle. The idea is to approximate the distribution of these distances by the observed distances between detections from the same frame. Indeed, the measured distances between detections in the same frame reflect the topology of detections originating either from particles or artifacts, but not from Brownian jumps since only one frame is considered. In the frame $k+1$ a detection that does not match any detection in the previous frame (false detection or newly appeared target) is expected to be localized according to the same topology as detections in frame k, but without regard to the position of the previous detections.

In practice, we measure the minimal distances between detections from the same frame and an empirical distribution is estimated by applying a kernel smoothing method [27]. This empirical pdf is used as the distribution of the minimal distances between detections in subsequent frames that do not correspond to Brownian jumps. By doing so the topology of the environment is naturally accounted for and some issues regarding the theoretical model of the distances pdf are avoided.

Having a statistical description of the distances between detections in subsequent frames, we are able to fit the diffusion model to the observed detections. To do so, we use an iterative procedure derived from the Expectation-Maximization (EM) algorithm [51]. The estimation algorithm is simplified in our case since only the parameter σ_b^2 has to be estimated for the Brownian jumps pdf, while the pdf of the distances between detections originating from different targets is already known. In practice we first evaluate the probability for each distance to originate from either of the statistical models, and then a new value of σ_b^2 is estimated thanks to the measured distances. As in the EM algorithm, the estimation of σ_b^2 is made with respect to the probabilities that a measurement originates from the two statistical models. The procedure is then iterated until convergence of the σ_b^2 estimate (less than one hundred iterations are generally required).

5.6.4 Experimental evaluation of the diffusion estimation procedure

In order to test the capability of the procedure to estimate accurately the diffusion process we have simulated the acquisition of 10 fluorescence images containing freely diffusing particles. The diffusion process is the same for every particle : its characteristic variance is chosen to be $\sigma_b^2 = 16\ pixels^2 frame^{-1}$, which is quite high compared to the movements generally observed in real images. The synthetic Gaussian noise which is added to the images has been adjusted to reach a PSNR of 5 and corresponds to a low SNR condition. In order to detect particles in images we have used the UDWT detection procedure presented in Chapter 4 : the second

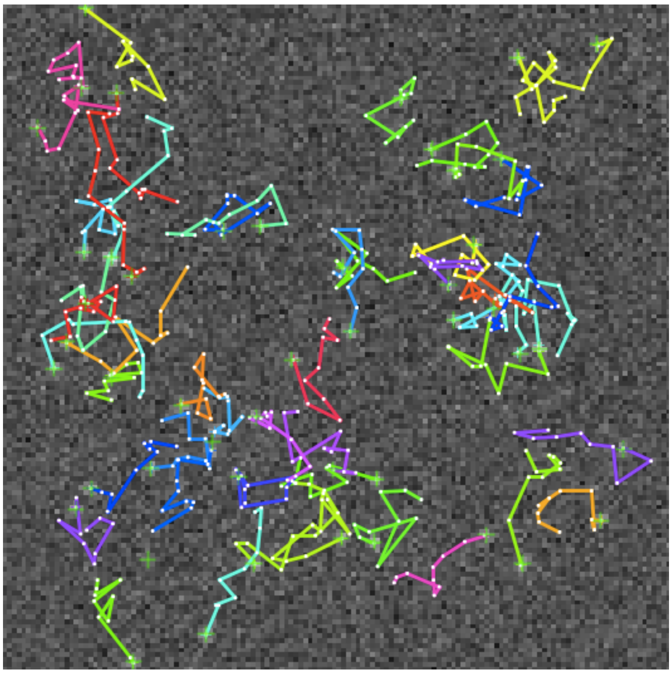


FIG. 5.6: Diffusion simulation in 2D : the Brownian motion of 50 particles is simulated for $\sigma_b^2 = 16\ pixel^2 frame^{-1}$ in 10 frames. Detections in the last frame are represented by green crosses.

and third wavelet scales are computed, thresholded according to the level of noise, and a noise free image is reconstructed. Particles are localized by binarizing this image and computing the center of mass of the segmented spots. Due to the high level of noise some particle positions are however missed by the detection procedure and some false detections are produced, which corresponds to realistic conditions. The corrupted detection set allows us to test the robustness of the estimation procedure against missed and false detections. We show in Figure 5.6 the last image of the sequence, and the corresponding tracks.

In order to estimate the diffusion coefficient without tracking particles we first compute for each detection in the whole sequence the distance to its closest detection in the same frame. The normalized histogram of these distances is given in Figure 5.7 a). We also show the curve corresponding to the theoretical pdf of the detections under the uniform spatial distribution assumption. The theoretical model provides a pretty good approximation of the observed distances, particularly for the decay of probabilities when the distance grows. However, the observed distances depart from the theoretical model for small distances. Moreover, very small distances which are predicted by the model are impossible to measure in practice due to the inability of the detection algorithm to discriminate closely spaced targets. In order to obtain a more accurate description of the pdf of minimal distances between detections the kernel smoothing method has also been used : the fitted distribution is superimposed to the normalized histogram in Figure 5.7 a).

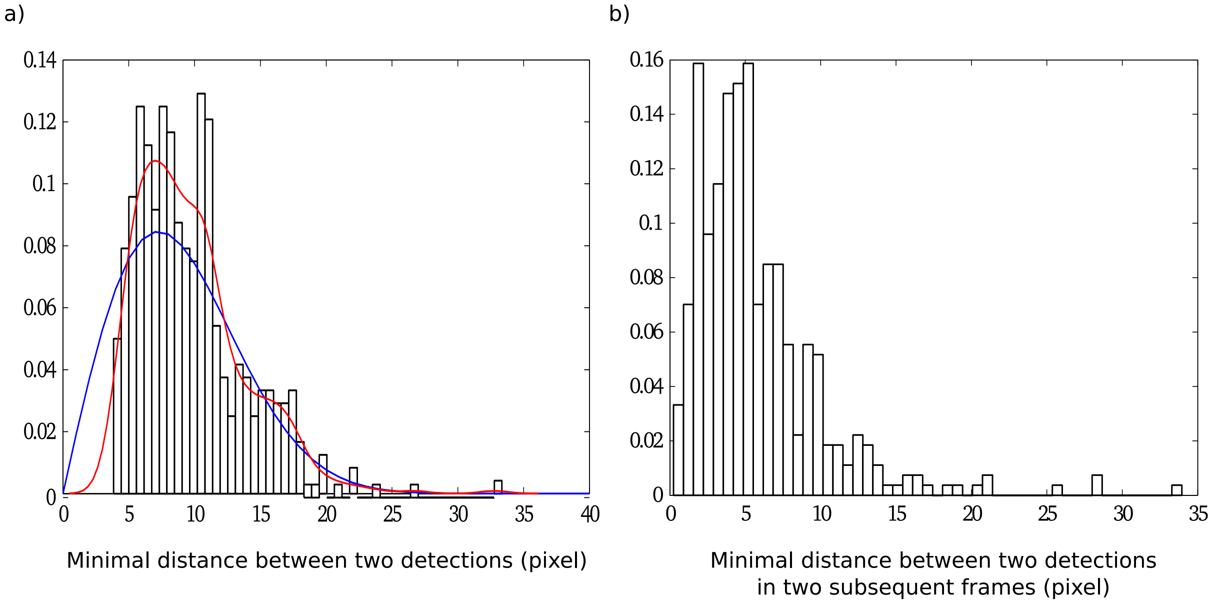


FIG. 5.7: Minimal distance histograms between detections in the images simulated for the diffusion estimation test. a) Normalized histogram of minimal distance between detections of the same frame. The theoretical distribution of the distance is plotted in blue, while the smoothed empirical distribution is represented by the red curve. b) Normalized histogram of distances for detections in subsequent frames.

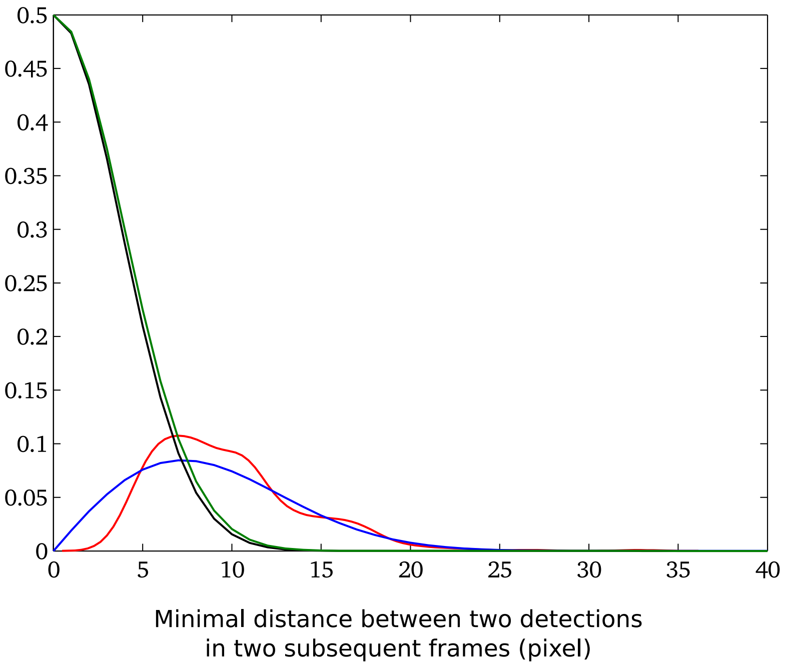


FIG. 5.8: The minimal distance distribution between detections in the same frame obtained by empirical estimation and theoretical computation are plotted in red and blue respectively. We plot in green the pdf of the Brownian jump length obtained by using the red pdf in the EM estimation procedure, while the black curve corresponds to the estimated distance pdf when the theoretical distribution is used.

We show in Figure 5.7 b) the normalized histogram of minimal distances between detections in subsequent frames, which have been computed by identifying for every detection of the sequence the closest detection in the next frame. We note that the histogram is mainly composed of two modes : one highly peaked mode near $d_{ij} = 2$, and one diffuse mode around $d_{ij} = 5$. When comparing the two histograms in Figure 5.7 we see that the mode around 5 is shared, but not the mode for $d_{ij} = 2$ which is specific to the right-hand side histogram. The shared mode therefore corresponds to the distances between detections that do not originate from the same particle, and we deduce that the mode corresponds to the Brownian jumps of particles. The presence of this well defined mode facilitates the estimation of the Brownian motion : using the modified EM algorithm we estimated σ_b^2 to be 15.65 when using the smoothed empirical distances distribution, which is very close to the true value 16. Using the theoretical pdf of distances we compute $\sigma_b^2 = 14.39$ which is a slightly less accurate estimation than with the empirical distribution. In practice one should thus use the empirical pdf estimation since it provides an accurate density estimation and accounts for detection topologies departing from uniformity. The theoretical model is however useful when the number of detections is not sufficient to compute a reliable estimation of the pdf of distances : it indeed does not depend on the observed distances.
The probabilistic distributions of the distances d_{ij} using the fitted σ_b^2 values are plotted in Figure 5.8. It shows that using either the empirical or the theoretical distribution results in a very similar estimation of the pdf of the distances. We also note that the distribution of Brownian jump length is broadly overlapping with the distribution of distances between detections which do not originate from the same target. Hence, using a standard EM algorithm estimating the two underlying probabilistic distributions would hardly provide as accurate results. We solve this issue by estimating one distribution either empirically or theoretically, regardless of the overlap between the two modes.

5.7 Summary

In this chapter we have described a statistical framework which is particularly suited to the task of tracking particles in fluorescence images because of its ability to model various biological systems and its ease of interpretation. We have defined as an objective function the probability of the tracks given the detections in the images. The tracks probability is expected to give relevant scores to tracking hypotheses since it is computed thanks to an exhaustive description of the tracking problem. As an example, the characteristics of the detection system are included through the statistical description of the false and missed detections, while the physical properties of the particles are accounted for through the statistical modeling of their motion. By doing so, changes of the real system properties find an echo in a straightforward way in the Bayesian model. The flexibility of the framework thus gives us the ability to design particle tracking algorithms which are appropriate to a variety of biological applications.

Chapter 6

Target perceivability modeling

6.1 Target perceivability concept

We derive next a model of target perceivability which is inspired from the work of Li [110, 111] and Musicki [124, 123] in other fields. The perceivability model we define allows the early detection of the end of a track and the initiation of tracks corresponding to real targets only. The approach is based on a HMM target existence description which we adapt to the tracking of particles in microscopy images. We also include the probabilistic model of perceivability inside the computation of the tracks likelihood $\mathcal{L}(\Theta)$ defined in Chapter 5, whereby the target tracking and track management are well posed in a common framework. In the proposed approach the track quality quantification is accounted for when computing association scores, so good quality tracks are favored and the robustness of the tracking procedure to false detections is consequently improved to a significant extent.

The perceivability of a target is defined as its capability to generate measurements [110]. We define a two states model of perceivability for particles :

- a particle is *perceivable* (state s^1) if it can be detected,
- a particle is *non perceivable* (state s^0) if it physically does not exist anymore or cannot be detected in images.

The perceivability state sequence is modeled as a first-order Markov chain, as shown in Figure 6.1, whereby the current perceivability state $s^j(k)$ depends on the previous perceivability state $s^i(k-1)$ but not on the earlier state sequence. We also define fixed transition probabilities

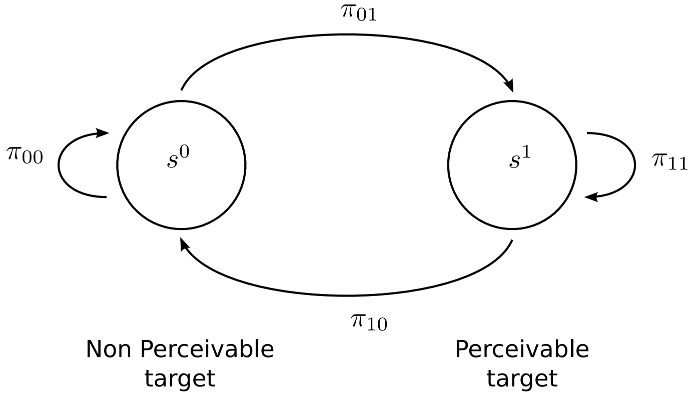


FIG. 6.1: First-order Markov chain for target perceivability.

between the state i and the state j, noted π_{ij}. The Markov assumption therefore results in computing the probability of being in the state s^i at time k as the probability of making the transitions from every state at time $k-1$, to the state s^i. We thus write the following recursion for the computation of the probability of the perceivability state s^i at a given time k :

$$p(s^i(k)) = \sum_{j\in\{0,1\}} \pi_{ji} p(s^j(k-1)). \tag{6.1}$$

We include the target perceivability concept in the Bayesian framework for tracking by writing the joint probability (θ^k, z^k) up to time k (Eq. 5.8) as follows[2] :

$$p(\theta^k, z^k) = \sum_{i\in\{0,1\}} p(\theta^k, z^k, s^i(k)), \tag{6.2}$$

hence explicitly accounting for the target perceivability. Using the perceivability concept in the estimation of the tracks likelihood favors perceivable tracks and thus increases the robustness of the tracking procedure to spurious measurements. Indeed, the association of a track with a false detection is expected to lower its probability of perceivability and a dummy target is expected to have a low probability of perceivability.
We can further factorize the track joint probability (Eq. 6.2) as :

$$p(\theta^k, z^k) = p(\theta^{k-1}, z^{k-1}) \sum_{i\in\{0,1\}} p(\theta(k), z(k), s^i(k)|\theta^{k-1}, z^{k-1}), \tag{6.3}$$

and then :

$$p(\theta^k, z^k) = p(\theta^{k-1}, z^{k-1}) \sum_{i\in\{0,1\}} \Big[p(\theta(k), z(k)|s^i(k), \theta^{k-1}, z^{k-1}) \cdot p(s^i(k)|\theta^{k-1}, z^{k-1})\Big]. \tag{6.4}$$

Equations 6.3 and 6.4 show that $p(\theta^k, z^k)$ can be evaluated by updating the joint state and measurement probability up to time $k-1$ with the current state and detection, hence allowing to compute $p(\theta^k, z^k)$ in an iterative way by estimating sequentially $p(\theta^r, z^r)$ with $r = 1..k$. The update process accounts for the possible perceivability states of the target by integrating out the different perceivability states. The update factor is indeed composed of two terms :

1. $\xi^i(k) \triangleq p(\theta(k), z(k)|s^i(k), \theta^{k-1}, z^{k-1})$ is the joint probability of the new target state $\theta(k)$ and measurement $z(k)$, given the perceivability state $s^i(k)$ at time k and the past measurements and states. The computation of this probability will be detailed in the next section.
2. $\lambda^i(k|k-1) \triangleq p(s^i(k)|\theta^{k-1}, z^{k-1})$ is the probability of the perceivability state s^i at time k when only past measurements and past states are known. We thus name it the *predicted probability* of the perceivability state s^i.

With these notations, and by also denoting $\chi^k \triangleq p(\theta^k, z^k)$, we can rewrite Equation 6.4 in a compact form :

$$\chi^k = \chi^{k-1} \sum_{i\in\{0,1\}} \xi^i(k)\lambda^i(k|k-1). \tag{6.5}$$

---

[2] For the sake of clarity we don't use bold fonts anymore for the measurements and states vectors in the remainder of the thesis. Hence z^k will refer to $\mathbf{z}^k$, x^k will refer to $\mathbf{x}^k$, etc.

Algorithm 2 Joint likelihood and perceivability sequential estimation.

```
Require: $z^l$, $\chi^1$, $\{\lambda^i(1|1)\}_{i\in\{0,1\}}$
  for $k = 2$ to $l$ do
    for $i = 0$ to 1 do
      $\lambda^i(k|k-1) \leftarrow \sum_{j\in\{0,1\}} \pi_{ji}\lambda^j(k-1|k-1)$ {Probability of perceivability prediction}
    end for
    for $i = 0$ to 1 do
      Compute $\xi^i(k)$ {See Section 6.2}
    end for
    $\chi^k \leftarrow \chi^{k-1} \sum_{i\in\{0,1\}} \xi^i(k)\lambda^i(k|k-1)$ {Update measurements and states joint probability}
    for $i = 0$ to 1 do
      $\lambda^i(k|k) \leftarrow \frac{\xi^i(k)\chi^{k-1}}{\chi^k}\lambda^i(k|k-1)$ {Update peceivability probability}
    end for
  end for
  return $\chi^l$
```

We also rewrite the predicted probability of perceivability at time k by integrating out the possible perceivability states at time $k-1$ and by using the Bayes' rule :

$$\lambda^i(k|k-1) = \sum_{j\in\{0,1\}} p(s^i(k)|s^j(k-1), \theta^{k-1}, z^{k-1})p(s^j(k-1)|\theta^{k-1}, z^{k-1}), \tag{6.6}$$

where $p(s^i(k)|s^j(k-1), \theta^{k-1}, z^{k-1}) = p(s^i(k)|s^j(k-1)) = \pi_{ji}$ from the Markov property of the perceivability chain and the conditional independence of the perceivability state on the target past states and measurements. The term $p(s^j(k-1)|\theta^{k-1}, z^{k-1})$ is the probability of the perceivability state s^j at time $k-1$ when the measurement and states are known up to this time. We decompose this probability as :

$$\begin{aligned} p(s^j(k-1)|\theta^{k-1}, z^{k-1}) &= \frac{p(s^j(k-1), \theta^{k-1}, z^{k-1})}{p(\theta^{k-1}, z^{k-1})}, \\ &= \frac{p(\theta(k-1), z(k-1)|s^j(k-1), \theta^{k-2}, z^{k-2})p(\theta^{k-2}, z^{k-2})p(s^j(k-1)|\theta^{k-2}, z^{k-2})}{p(\theta^{k-1}, z^{k-1})}, \end{aligned} \tag{6.7}$$

which can be rewritten with the notations introduced above as :

$$p(s^j(k-1)|\theta^{k-1}, z^{k-1}) = \frac{\xi^j(k-1)\chi^{k-2}}{\chi^{k-1}}\lambda^j(k-1|k-2). \tag{6.8}$$

Equation 6.8 shows that the probability of the perceivability state s^j at a given time $k-1$ is computed as an update of its predicted probability of perceivability $\lambda^j(k-1|k-2)$ with the additional data found in time $k-1$. We therefore adopt the notation $\lambda^i(k|k) \triangleq p(s^i(k)|\theta^k, z^k)$ for any arbitrary state s^i and time k, and we rewrite Equation 6.6 as :

$$\lambda^i(k|k-1) = \sum_{j\in\{0,1\}} \pi_{ji}\lambda^j(k-1|k-1). \tag{6.9}$$

By combining Equation 6.5 with Equations 6.8 and 6.9 we can now design a sequential process for the estimation of χ^k : it consists in alternating the estimation of χ^r and that of $\{\lambda^i(r|r-1)\}_{i\in\{0,1\}}$ with $r = 1..k$. Algorithm 2 describes the complete estimation scheme which consists in a loop with three main steps : (1) the probability of each perceivability state is first predicted, (2) a new measurement and state are then considered, which leads to the update of the joint probability of the sequence of states and measurements, and (3) the perceivability probabilities are updated in consequence. Each loop appends a new measurement and a new state until the whole track is processed, hence making the probability estimation compatible with standard tracking procedures which sequentially link detections over time. The recursive property of the estimation algorithm is also very desirable since all the previous data do not have to be re-processed whenever a new measurement is received.

We explicit next the computation of the probability of the current measurement and state conditioned to the perceivability state : $\xi^i(k)$, which can take four forms depending on the target perceivability and on whether the detection is real or virtual.

6.2 Measurement and state joint probability estimation

6.2.1 Perceivable target

The state estimation procedure introduced in Chapter 5 can be used to compute the joint probability $\xi^1(k)$ of a perceivable target since the target is implicitly assumed perceivable in this model. Applying the Bayes' rule we write :

$$\xi^1(k) = p(z(k)|s^1(k), \theta^k, z^{k-1})p(\theta(k)|s^1(k), \theta^{k-1}, z^{k-1}). \tag{6.10}$$

where the measurement probability $p(z(k)|s^1(k), \theta^k, z^{k-1})$ depends on the type of the detection $z(k)$. Indeed, if $z(k)$ is a real detection it is assumed to originate from the target intensity, which implies that the target has been detected and that the detection has fallen into the search gate. In this case the conditional detection probability is computed as : $p(z(k)|s^1(k), \theta^k, z^{k-1}) = P_D P_G g(z(k)|\theta^k, z^{k-1})$ where P_D is the probability of detecting the target when it is perceivable, P_G is the probability that the corresponding detection falls into the search gate, and $g(z(k)|\theta^k, z^{k-1})$ is the detection probability provided by the measurement model in the state-space (Eq. 3.10).

In contrast, if the detection is virtual we consider either that the target has not been detected, or have left the search gate. We also denote $g(\theta(k)|\theta^{k-1}, z^{k-1}) \triangleq p(\theta(k)|s^1(k), \theta^{k-1}, z^{k-1})$ the state evolution probability when the target is perceivable, which results in :

$$\xi^1(k) = \begin{cases} P_D P_G g(z(k)|\theta^k, z^{k-1}) g(\theta(k)|\theta^{k-1}, z^{k-1}) & \text{if } z(k) \text{ is real,} \\ 1 - P_D P_G & \text{else,} \end{cases} \tag{6.11}$$

for the joint state and measurement probability when the target is perceivable.

6.2.2 Non perceivable target

A non perceivable target does not generate any detection since it cannot be detected by definition of the perceivability concept. Hence, if the measurement associated to a non perceivable target is real we consider it as being a false detection. On the other hand, associating a virtual measurement to a non existing target is valid and we do not have to consider this association since both the measurement and the track do not exist. The computation of $\xi^0(k)$ is therefore achieved with respect to the type of detection :

$$\xi^0(k) = \begin{cases} p_{fd}(z(k), k) & \text{if } z(k) \text{ is real,} \\ 1 & \text{else,} \end{cases} \tag{6.12}$$

Detection type	Perceivable target	Non-perceivable target
Real	$\xi^1(k) = P_D P_G g(z(k)\vert\theta^k, z^{k-1}) g(\theta(k)\vert\theta^{k-1}, z^{k-1})$	$\xi^0(k) = P_{FD}(\mathbf{z}(k), k)$
Virtual	$\xi^1(k) = 1 - P_D P_G$	$\xi^0(k) = 1$

TAB. 6.1: Computation of $\xi^i(k)$ as a function of the state $s^i(k)$ and the type of the associated measurement $z(k)$. The measurement is said to *real* if it is provided by the particle detection procedure, while a *virtual* measurement is provided by the motion filter to compensate for the lack of associated detection to the target.

where $p_{fd}(z(k), k)$ is the probability that the detection $z(k)$ originates from clutter. The probability $p_{fd}(z(k), k)$ has to be chosen with care since we aim here at providing a model that can be incorporated in the Bayesian framework presented in Chapter 5. In this framework non perceivable targets are not accounted for, so, we have to tune the perceivability model such as detections associated to a non perceivable target ($\lambda^1(k|k) = 0$) participate to the computation of the tracks likelihood (Eq. 5.8) in the same way as a non assigned measurement. In the standard framework non assigned detections are considered as false detections and their probability is computed by Equation 5.31. We therefore deduce from Equations 6.5 and 5.31 that the following equalities have to be satisfied :

$$\lim_{\lambda^1(k|k-1)\to 0} \chi^k = \chi^{k-1} P_{FD}(\mathbf{z}(k), k), \tag{6.13}$$

$$\Rightarrow \chi^{k-1}\xi^0(k) = \chi^{k-1} P_{FD}(\mathbf{z}(k), k), \tag{6.14}$$

when the measurement $z(k)$ is real. This last statement leads us to set $p_{fd}(z(k), k) = P_{FD}(\mathbf{z}(k), k)$. By doing so the contribution to the objective score function is the same when a detection is not assigned to a track and when it is assigned to a non perceivable track. Non assigned detections can therefore be viewed as belonging to a dummy target which is not perceivable with probability 1. By choosing such settings the perceivability concept is perfectly embedded within the Bayesian framework. We therefore expect a reliable behavior of the objective score function.

We summarize in Table 6.1 the different manners to compute the joint probability $\xi^i(k)$ in function of the type of detection (virtual or real) and the perceivability the state of the target (perceivable and non-perceivable).

6.3 Initial probability of perceivability

Providing the initial probability of perceivability[3] of target $\{\lambda^i(1|1)\}_{i\in\{0,1\}}$ is required to initiate the recursive estimation procedure described by the Algorithm 2. When a track is created from a measurement the only prior information we have on whether the detection originates from a target and not from the noise is the prior distribution of target appearance events and the prior distribution of the false detections. We consequently base the definition of the initial probability of perceivability on the probability of these two kind of events.

In Chapter 5 we have assumed for each voxel $\mathbf{x} = \{x, y, z\}$ a probability $P_{FD}(\mathbf{x}, k)$ to generate a false measurement, which is position dependent (Eq. 5.31). We have also assumed that in each frame a new target can appear at a given pixel with the probability $P_{NT}(\mathbf{x}, k)$ which is also

---

[3] The perceivability estimation procedure is started here in the frame $k = 1$, without loss of generality for initiating the process at other time points.

position and time dependent (Eq. 5.26). When the probability of generating a false detection $P_{FD}(\mathbf{x}, k)$ is very small compared to the probability that a new target appears, we are almost sure that a detection at the location $\mathbf{x}$ which is non assigned to a track can only originate from a perceivable target. Hence, intuitively $\lambda^1(1|1)$ should be close to one. Conversely, if the probability of target appearance is much smaller than the probability of false detection : $P_{FD}(\mathbf{x}, k) \gg P_{NT}(\mathbf{x}, k)$, we are almost sure that such a detection originates from noise. These two statements lead us to write two requirements for the design of the initial probability of perceivability :

$$(A) : \lim_{P_{NT}(\mathbf{x},k)/P_{FD}(\mathbf{x},k)\to\infty} \lambda^1(1|1) = 1, \tag{6.15}$$

$$(B) : \lim_{P_{NT}(\mathbf{x},k)/P_{FD}(\mathbf{x},k)\to 0} \lambda^1(1|1) = 0. \tag{6.16}$$

Based on requirements (A) and (B) we propose to chose the initial probability of perceivability according to :

$$\lambda^0(1|1) = \frac{P_{FD}(\mathbf{x}, k)}{P_{FD}(\mathbf{x}, k) + P_{NT}(\mathbf{x}, k)} \ , \ \lambda^1(1|1) = \frac{P_{NT}(\mathbf{x}, k)}{P_{FD}(\mathbf{x}, k) + P_{NT}(\mathbf{x}, k)}, \tag{6.17}$$

and we compute the first state and measurement probability as : $\chi^1 = \sum_{i\in\{0,1\}} \xi^i(1)\lambda^1(1|1)$. The estimation loop defined in Algorithm 2 is then applied to the remaining measurements and states.

The design of the initial probability of perceivability proposed in Equation 6.17 is capable of modeling complex biological systems since $\lambda^1(1|1)$ can be both time and position varying. The value of $\lambda^1(1|1)$ is indeed determined by the probabilities $P_{NT}(\mathbf{x}, k)$ and $P_{FD}(\mathbf{x}, k)$ which are both time and position dependent. The position dependent model of initial perceivability hence easily accounts for topological specificity by attributing different probabilities of perceivability to candidate tracks lying in different regions of the stack. As an example, the appearance of new targets may be more frequent in some parts of the image due to the topology of the environment. When the probabilities of false detection and new target appearance are both constant in space and time the initial state probabilities are written in a more compact form :

$$\lambda^0(1|1) = \frac{P_{FD}}{P_{FD} + P_{NT}} \ , \ \lambda^1(1|1) = \frac{P_{NT}}{P_{FD} + P_{NT}}, \tag{6.18}$$

and are also constant in space and time.

6.4 Track life decisions

A non perceivable target (state s^0) may have physically disappeared, bleached to a low level of intensity or left the surveillance volume. Such a target will not produce any measurement in future frames, so, we would like to stop the tracking process by ending the corresponding track. In contrast, a perceivable target (state s^1) generates photons that are visible in the image sequence, so tracking it over time is wanted. These two remarks motivate the use of the perceivability concept for track life decision : tracks should be created for perceivable targets only, and tracks should be terminated as soon as targets become non perceivable.

Contrary to most tracking algorithms we do not create automatically a track for each non assigned detection, but we prefer instead to distinguish two sets of tracks : *confirmed* tracks and *candidate* tracks. Candidate tracks correspond to targets whose probability of perceivability is low. In this case the risk that the detections associated to them originate from noise is high, hence we reject these tracks until their perceivability is high enough. We thus define a confirmation threshold λ_c on the target probability of perceivability :

$$\theta \text{ is confirmed at time } k \Leftrightarrow \exists k' \leq k \text{ s.t. } \lambda^1(k'|k') > \lambda_c, \tag{6.19}$$

which means that a candidate track is confirmed at a given time only if its probability of perceivability has exceeded λ_c in past frames. The tracking algorithm proposed in Chapter 7 uses only the confirmed tracks for solving the association problem between tracks and detections whereby its size is significantly lowered.

Similarly as for track creation we use the probability of perceivability for deciding to stop the tracking process : at time k if the predicted probability of the target falls beyond the termination threshold λ_t : $\lambda^1(k+1|k) < \lambda_t$, we stop the track since it is very probable that the target will not be perceivable again in future frames.

The design of the threshold λ_c is a delicate task, as we should confirm the highest number of tracks corresponding to perceivable targets, while confirming very few tracks created from spurious detections. λ_t should also be tuned to stop tracks as soon as targets become non perceivable, while still allowing some low probability associations. Thresholds confirming only perceivable targets and stopping tracks as soon as targets disappear are however difficult to define in practice due to the presence of numerous false detections and some movements with a low probability. In Chapter 7 we will thus design a tracking algorithm that is capable of accommodating numerous wrongly confirmed tracks.

6.5 State transition parameters

Discriminating a new particle from a non-perceivable target becoming perceivable again after a long time interval is barely feasible for fluorescent objects which are almost alike. We therefore assume that non-perceivable targets cannot become perceivable again, and we set $\pi_{01} = 0$. Linking partial tracks originating from the same particle can still be done in a post-processing step as done in [25] and in [94]. The other state transition parameters are not *a priori* known for most biological applications, but fixing them beforehand is required.

The perceivability state model can be seen as a Bernouilli process : between subsequent frames one perceivable particle stays perceivable with probability π_{10} and can become non perceivable with probability π_{10}. As soon as the target becomes non perceivable the process is stopped. For a particle perceivable at frame t the prior pdf of switching to the non perceivable state at frame $t+\delta t$ is given by the geometric distribution : $Pr(X = t+\delta t) = (1-\pi_{10})^{\delta t}\pi_{10}$. The mean delay before switching to the state s^0 is therefore given by π_{10}^{-1}. We propose to fix the transition parameter π_{10} in such a manner that the expected track length given by the perceivability model matches the prior expected track length : l_m. To do so we set :

$$\pi_{01} = \frac{1}{l_m}. \tag{6.20}$$

The mean track length l_m is an easily interpretable parameter which can be evaluated with a short visual inspection of the data set. We generally fix $\pi_{01} = 0.05$, hence expecting 20 frames length tracks. Fixing roughly π_{01} to an approximated value still provides the desired behavior of the perceivability estimation procedure since it also relies on the observation model, and not on the prior transition model solely. We will indeed see with some experiments that slightly different π_{01} settings do not induce dramatic changes of the estimated probability of perceivability. The estimation procedure is thus robust to slight variations of the transition settings.

The two remaining transition parameters are simply fixed as : $\pi_{11} = 1 - \pi_{10}$ and $\pi_{00} = 1$.

6.6 Simulations and experiments

Incorporating the perceivability estimation procedure into the Bayesian framework presented in Chapter 5 results in a model capable of describing exhaustively the tracking problem. The complexity of this model is however high and various statistical tools and tuning parameters are involved. We test next the influence of varying the perceivability model parameters to check if the model behaves in a coherent manner, despite its complexity.
In order to check the accurate behavior of the perceivability estimation procedure we have simulated the diffusion of a particle in a 3D environment. The particle freely diffuses during 20 frames, and then disappears from the images. The coefficient characterizing the diffusion process is set to : $\sigma_b^2 = 9 \ pixels^2 frames^{-1}$. We show in Figure 6.2 some 2D and 3D views of the resulting trajectory, which highlights that between frames 5 and 6 a large particle jump occurs. This low probable movement will serve to analyze the behavior of the perceivability model when events with a low probability occur. In order to check the behavior of the perceivability model when objects of interest are missed by the detection procedure we will also eventually remove the measurement from the frame 10.
We investigate next the role of three parameters : the false detection probability : P_{FD}, which is assumed constant over time and space, the probability of detecting the particle P_d, and the transition probability π_{10} between the perceivable and non perceivable states. To do so, we vary each parameter one by one, other parameters being fixed. The default model settings are the following : $P_{FD} = 0.0001$, $P_d = 0.85$, $\pi_{10} = 0.05$, $P_{NT} = 0.00001$. We do not investigate here the role of the probability P_{NT} of new target since it only influences the initial probability of perceivability, nor do we study the probability of falling into the gate P_G since its role is similar to the role of P_D in the estimation procedure.

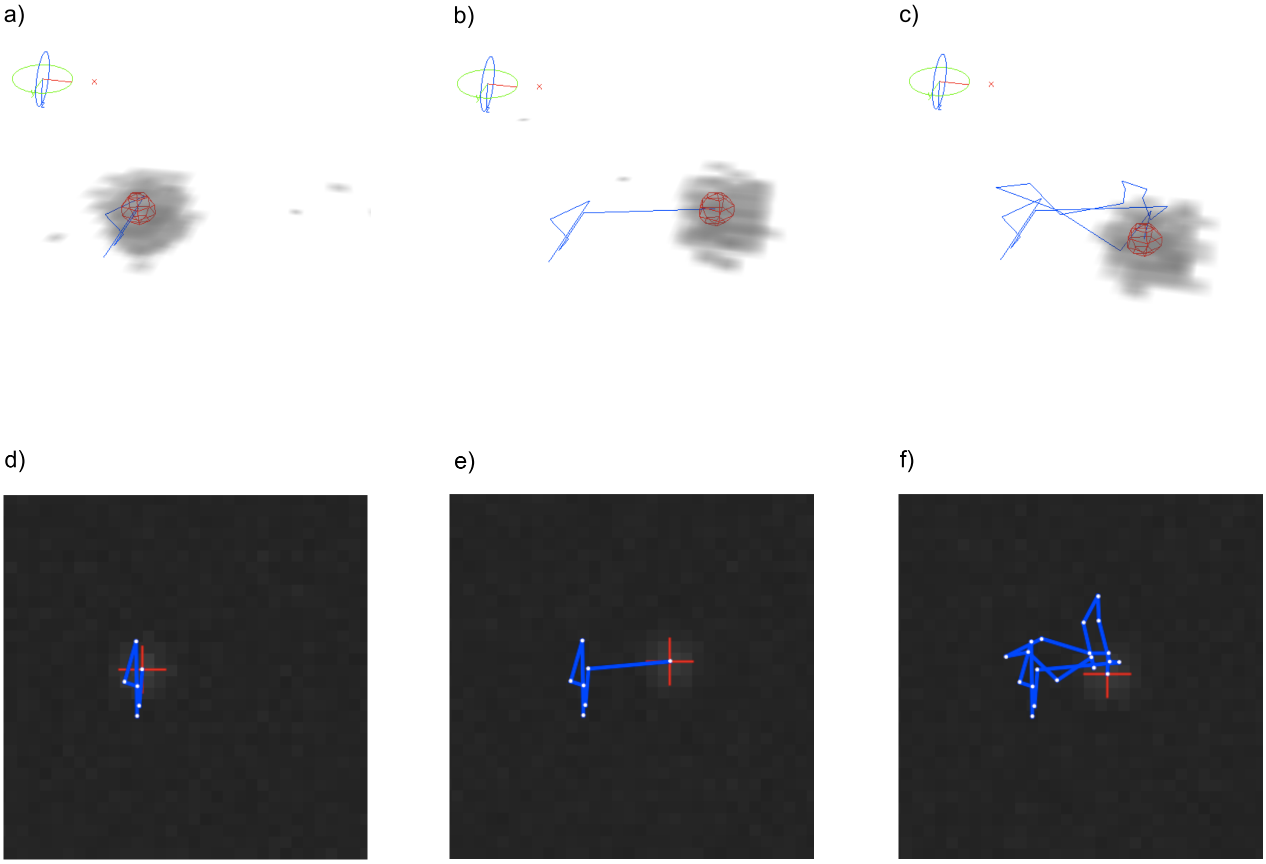


FIG. 6.2: Simulation of the particle diffusion for the evaluation of perceivability model. We simulate the diffusion of a particle in a 3D environment for 20 frames. a), b) an c) are the 3D stacks corresponding to frames 5, 6 and 20, respectively. d) to f) are the 2D maximum intensity projections of the same stacks. We show in blue the track corresponding to the particle movements. Between frames 5 and 6 a long jump with low probability occurs.

6.6.1 Varying the false detection probability

In order to check the influence of the false detection probability on the perceivability estimation we have used the default settings with five different values of P_{FD} : 0.0001, 0.0002, 0.0003, 0.0004 and 0.0005. We show in Figure 6.3 the resulting perceivability probabilities for the diffusing particle.

The probability of false detection plays an important role in the estimation of the perceivability state : the greater the density of false detections, the less probable the perceivable state. The density probability has a strong impact on the growth rate of the probability of perceivability. We show indeed that when P_{FD} is high ($P_{FD} = 0.0005$) reaching a probability of perceivability superior to 80% requires 13 subsequent measurements, which is very long. This slow growth rate is mainly due to the fact that the false detection model gives high probability to the measurements when P_{FD} is large. The target model thus does not provide a significantly better explanation of the data in this case. In contrast, the perceivablity values grow fast when the spurious detection model gives very low probability values to the measurements. As an example, we evaluate $\lambda^1(3|3) = 0.85$ and $\lambda^1(4|4) = 0.99$ when $P_{FD} = 0.0001$.

The perceivability curves all show a decrease at frame 7 after the unlikely movement, the extent of the decrease of the probability being linked to the value of P_{FD}. This behavior shows that the perceivability model indeed accounts for the statistical model of motion : associating detections to a target does not necessarily increases its probability of perceivability. Associating particles that agree with the model of motion of the target is necessary to improve the perceivability since the false detection model can give high probabilities to the measurement when P_{FD} is large. These observations highlight the key importance of the accuracy of the motion and false detection models for the reliable estimation of the target perceivability probabilities.

Interestingly, we show in Figure 6.3 that the decrease of the probability of perceivability is not linked to P_{FD} when no more measurements are associated to a track. Indeed, the false detection probability P_{FD} is no longer involved in the estimation of probability of perceivability when the detection is virtual (Eq. 6.11 and Eq. 6.12). A similar statement holds in Figure 6.3 b) for the missed detection at frame 10 : the main difference between the various P_{FD} settings is the rapidity of the recovery after the sudden probability decrease resulting from the association with a virtual measurement.

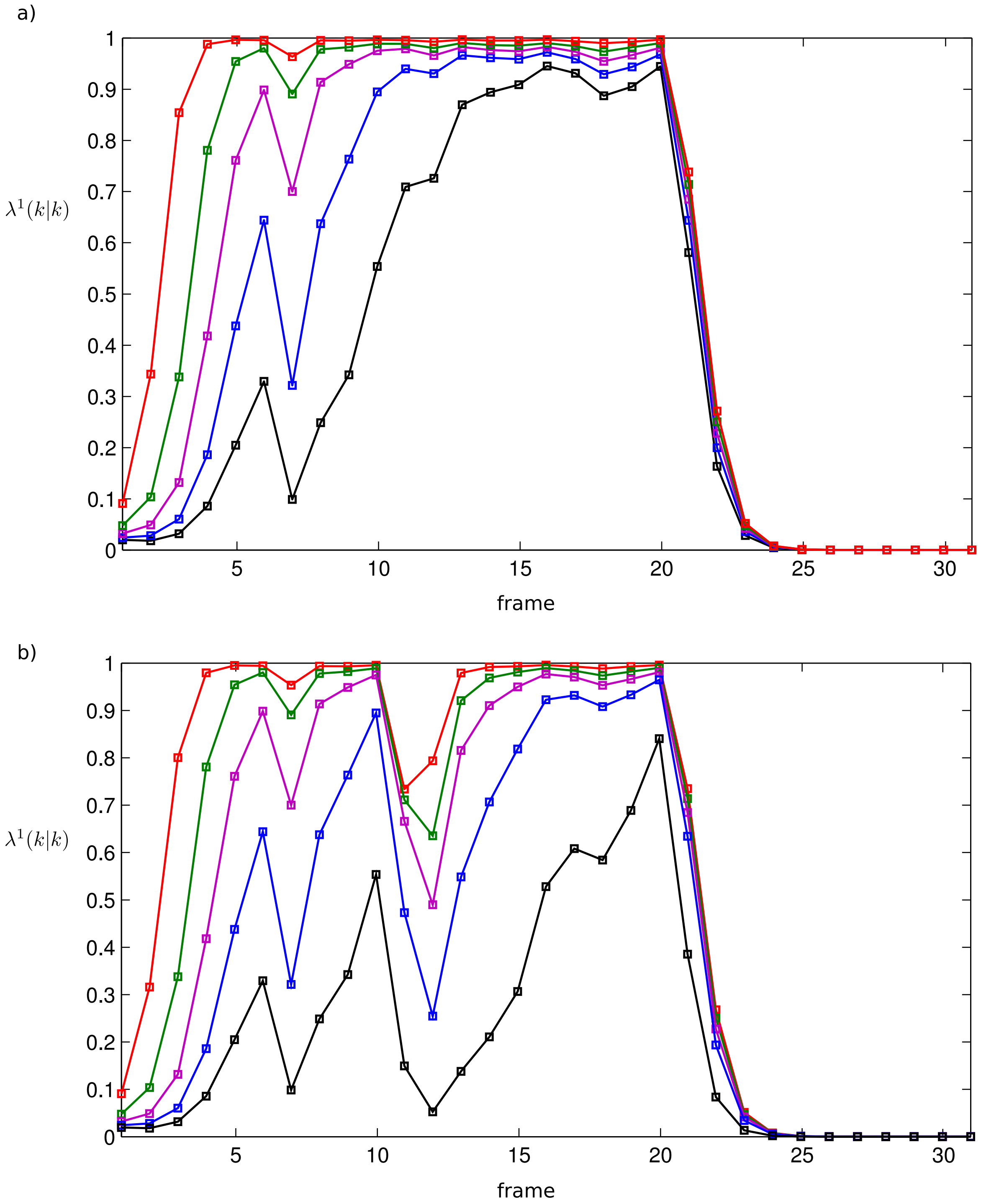


FIG. 6.3: Perceivability estimation with different values of the probability of false detection : $P_{FD} = 0.0001$ (red curve), 0.0002 (green), 0.0003 (purple), 0.0004 (blue) and 0.0005 (black). Figures a) and b) correspond to the same trajectory, but a detection has been removed at time 10 in the second case (b) to simulate a missed detection.

6.6.2 The role of the probability of detection

We have estimated the perceivability of the diffusing particle for six different settings of the probability of detecting the target : $P_D = 1$, 0.9, 0.8, 0.7, 0.6 and 0.5. The resulting curves are given in Figure 6.4. We observe very different features than when varying P_{FD} :

- The initial growth of the perceivability is influenced by P_D only to a limited extent. The same observation holds for the association with the unlikely detection in frame 7.
- The extent of the decrease of probability when no more detections are associated to the track is strongly influenced by the value of P_D : 7 frames are required to drop below a probability of 0.1 when $P_D = 0.5$, while just two associations with virtual detections result in a probability of perceivability of 0.002 when $P_D = 1$.

We show in Figure 6.4 b) that a growing P_D value has opposite effects when the track is associated to a real or to a virtual detection. Indeed, the probability of perceivability drops to a very low value when the target is not associated to a real detection and when P_D is large. It is indeed low probable that the detection procedure has missed the target. On the contrary, associating the target to a real detection results in higher values of the probability of perceivability when P_D is high than when it is low. The opposite effects of the value of P_D come from the negative and positive impacts it has on the value of $\xi^1(k)$ depending on the nature of the detection, as shown in Equation 6.11.

The general trend observed with the diffusion experiment is the faster adaptation of the values of perceivability probability when P_D is high. Real detections indeed result in a fast growth, and associations with virtual detections result in a fast decrease. This tendency is easily interpretable : when detecting the target is uncertain only a limited confidence is given to the measurements (since they can originate from the noise), and to the missed detections which can be due to the failure of the detection procedure. On the contrary, a high probability of detections results in a high confidence in the observed data.

6.6.3 State transition parameter influence

We have investigated the influence of the state transition probability π_{10} on the perceivability evaluation by applying the estimation procedure with 5 different settings : $\pi_{10} = 0.05$, 0.15, 0.25, 0.35 and 0.45. The resulting perceivability curves are shown in Figure 6.5. As expected, we observe that the higher the probability of switching to a non perceivable state, the lower is $\lambda^1(k|k)$, the probability of being perceivable. We indeed note that the base line of the perceivability curves decreases with an increase of π_{10}, even if the associations of the track with measurements are probable. The maximum achievable probability of perceivability is thus linked to the value of π_{10}.

In Figure 6.5 we also show that the decrease extent of $\lambda^1(k|k)$ varies with π_{10} when unlikely movements occur and when a detection is missed. This property is coherent with the model of target perceivability : an event which has a low probability according to the perceivable target model induces a significant decrease of the probability of being perceivable when the transition from the state s^1 to the state s^0 is probable. On the other hand, if the transition to the non perceivable state has a low probability, the observation is more probable for the perceivable state than switching to the non perceivable state and generating this measurement.

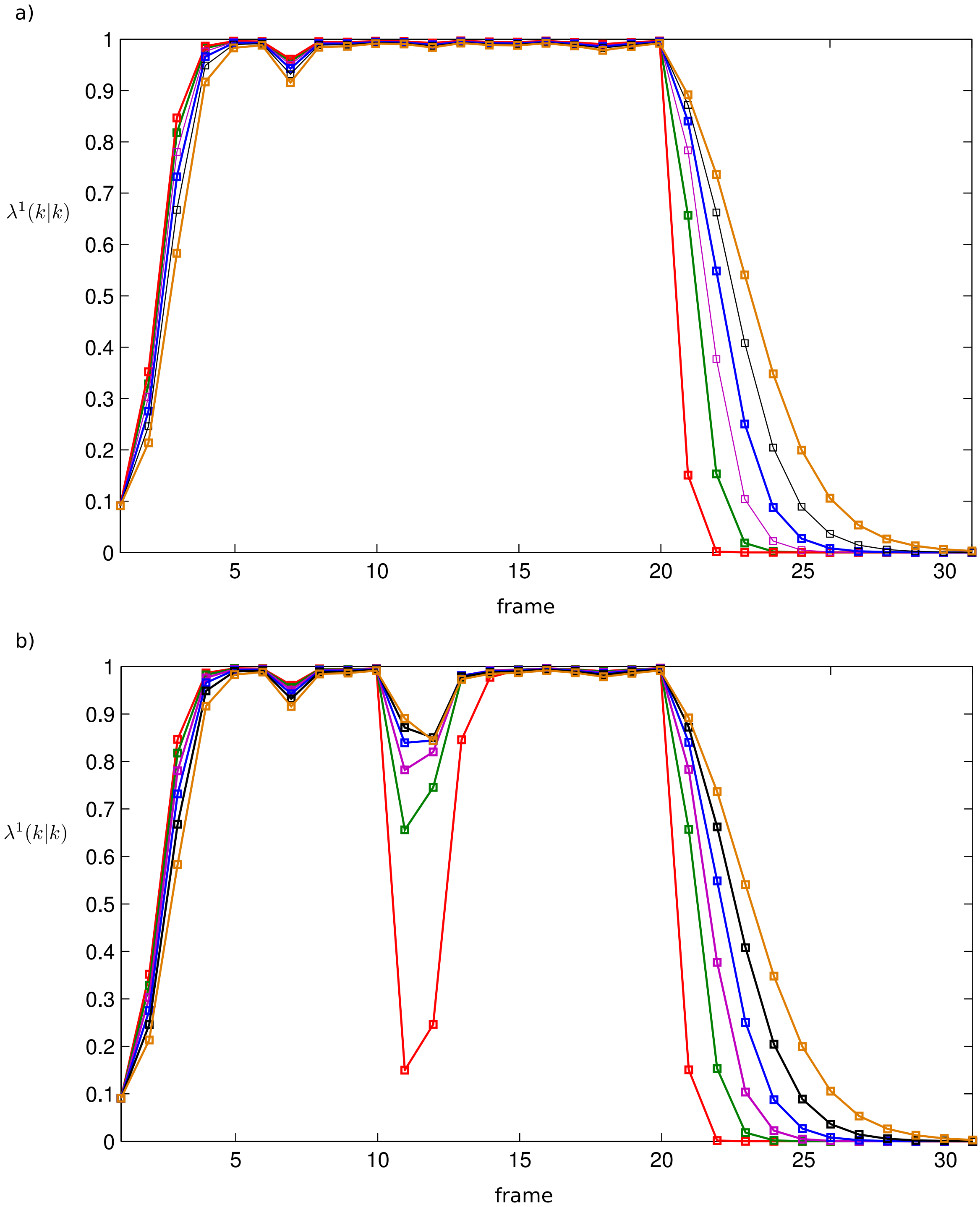
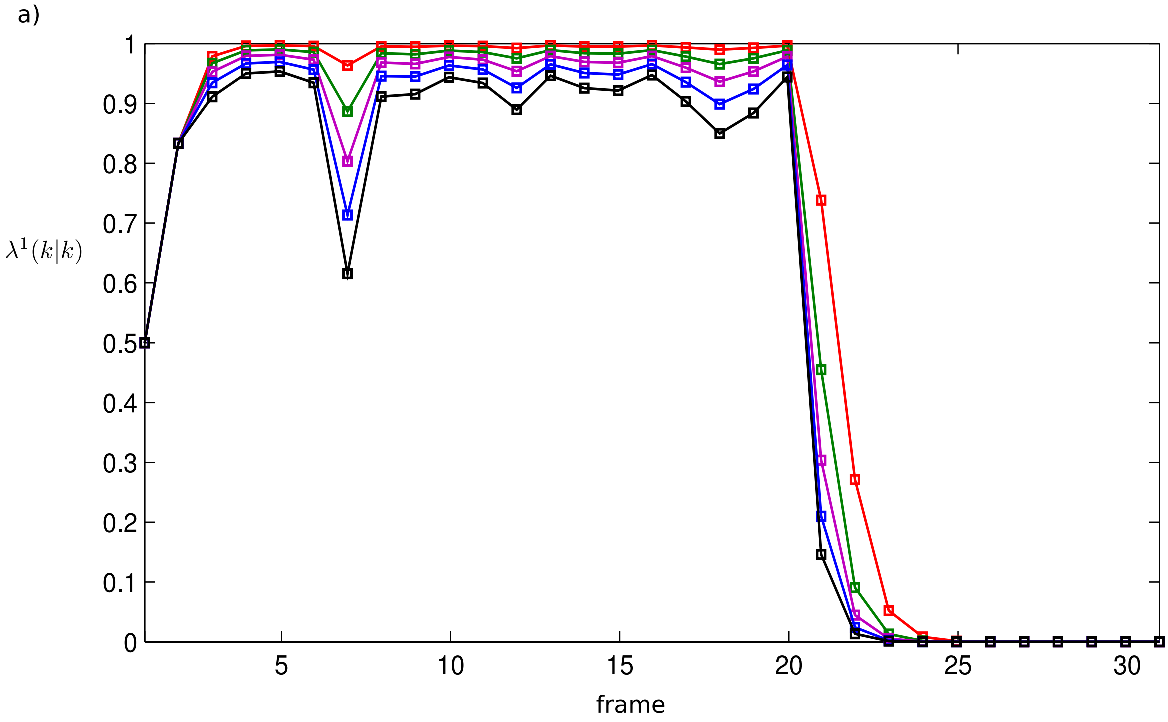


Fig. 6.4: Perceivability estimation with different values of the probability of detecting the particle : $P_D = 1$ (red curve), 0.9 (green), 0.8 (purple), 0.7 (blue), 0.6 (black) and 0.5 (orange). Figures a) and b) correspond to the same trajectory, but a detection has been removed at time 10 in the second case (b) to simulate a missed detection.



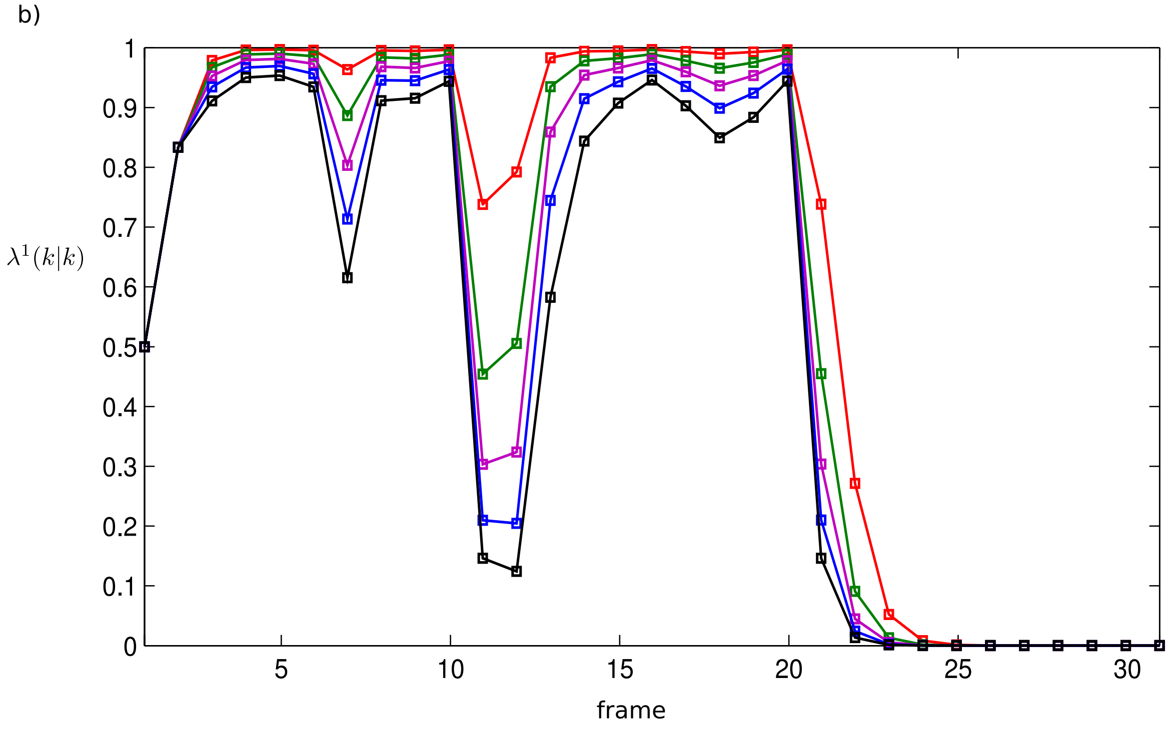


FIG. 6.5: Perceivability estimation with different values of the probability of transition between the perceivable state and the non perceivable state : $\pi_{10} = 0.05$ (red curve), 0.15 (green), 0.25 (purple), 0.35 (blue) and 0.45 (black). Figures a) and b) correspond to the same trajectory, but a detection has been removed at time 10 in the second case (b) to simulate a missed detection.

6.6.4 Experiments summary

The three experiments we have conducted for the estimation of the perceivability probability of a particle diffusing in a 3D environments highlight the intuitive and easily interpretable behavior of this yet complex model :

- a high probability of false detection lowers the probability of perceivability since measurements are likely to originate from artifacts,
- a high probability of detecting the particle strongly penalizes the missed detection event,
- the probability of being perceivable is lowered by a high π_{10} value since switching to the non perceivable state is probable.

We have verified that events such as movements with a low probability or missed detections are appropriately accounted for by the model. Moreover, as shown in Figures 6.3, 6.4 and 6.5, varying the model parameters induces smooth variations of the estimated $\lambda^1(k|k)$ values, without any sudden and unpredictable changes. The parameters defining the target perceivability have thus to be set to appropriate values, but a very accurate tuning of the model is not required : rough estimations of the parameters are sufficient to obtain the appropriate behavior of the estimation procedure. These desirable properties make the perceivability estimation algorithm suitable to statistical models of particle tracking.

6.7 Conclusion

In this chapter we have proposed a statistical model of target perceivability which allows us to estimate the probability that a target will produce measurements in future frames. This model is based on a HMM with two states : perceivable and non perceivable target, which are characterized by their statistics of generating both real and virtual measurements. We provide a recursive procedure to estimate the probability that one target is perceivable by exploiting these statistics when analyzing the sequence of measurements which are associated to it. We have experimentally validated this estimation procedure with synthetic data, and we have proved the desirable behavior of the procedure with different settings.

The probability of perceivability is an accurate indicator of target existence which we thus use as a statistical criterion for creating and stopping tracks. We have also shown that the perceivability model can be incorporated in the Bayesian framework for tracking by accounting for the possible perceivability states when computing the tracks likelihood. By doing so, tracks corresponding to perceivable targets are favored compared to tracks of particles which have probably disappeared. We therefore expect to obtain more reliable likelihood values when tracking multiple targets in cluttered conditions.

Chapter 7

Enhanced Multiple Hypothesis Tracking

In this chapter we present a new formulation of the Multiple Hypothesis Tracking (MHT) [138] that does serve two purposes : first we develop an operative implementation of the MHT for numerous targets in a dense clutter environment, and second, it optimally optimizes a statistical score function that includes the target perceivability concept, as well as dedicated models of motion for biological particle tracking.

7.1 Multiple Hypothesis Tracking

The MHT has been proposed almost thirty years ago to cope with the limitations of instantaneous tracking. The principle of the MHT is to delay the association step to a later time when the decision is made easier by the knowledge of future frames. In practice it relies on building many possible associations between tracks and detections for a number of subsequent frames and comparing them. Since it takes advantage of temporal information, the MHT is generally accepted as the method of reference for solving the data association problem in modern multiple target tracking systems (see [20] and references therein). The main benefits of the MHT technique lie in its unique capability of :

- automatically initiating tracks as targets appear,
- automatically ending tracks as targets disappear,
- continuing tracks when targets transiently disappear,
- explicitly modeling spurious measurements presence,
- explicitly ensuring that a track associates to a single measurement, and conversely.

Unfortunately these features come at the price of an exponentially growing computational and memory requirements with the number of accounted frames and the number of detections, which has prevented its use in practical biological applications.

7.1.1 Reid's algorithm

The first MHT implementation by Reid [138] (also coined *Reid's algorithm*) is said *hypothesis-oriented* since it consists in building few tracking hypotheses and extend them through time. The set of feasible tracking hypotheses up to time k is denoted Ω^k. A tracking hypothesis $\Theta^k \in \Omega^k$ can be extended in frame $k+1$ in numerous ways depending on the possible associations of tracks in Θ^k with the measurements in $Z(k+1)$. The sequence of feasible tracking hypotheses

$\{\Omega^k\}_{k=1..l}$ can therefore be represented with a tree where a node at the k^{th} level Θ^k gives rise to a number of nodes at the level $k+1$ depending on the associations at time $k+1$. The Reid's algorithm relies on iteratively extending the hypothesis tree with additional levels and pruning and merging some branches to avoid an exponential explosion of the number of leaves in the trees. Hypotheses pruning is based on a association likelihood thresholding criterion whereby only high probability hypotheses are kept, while hypotheses merging is based on a similarity criterion between hypotheses which allows one to group hypotheses with similar consequences.

7.1.2 Modern implementations of the MHT

The advantages of the Reid's algorithm over instantaneous tracking have lead to the development of a number of MHT algorithms over past decades (see [18] for an up-to-date review). More specifically, the use of the Murty's algorithm [122] to create efficiently the m-best hypotheses has significantly decreased the computational cost of the hypotheses generation step [47]. Indeed, using only the m-best hypotheses at each frame avoids the generation of many low probability hypotheses which would considerably slow down the algorithm. Doing so is nevertheless a heuristic approach to MHT with an accuracy which is difficult to assess.

A number of MHT implementations follow the hypothesis-oriented approach of Reid, however an alternative approach, coined *track-oriented*, has been introduced by Kurien [107] and refined over years [18]. The main ingredient of track-oriented MHT algorithms is to maintain one tree of association for all tracks instead of numerous trees for global association hypotheses. A track tree is composed of the possible sequences of associations of a track with detections. At each frame global association hypotheses are generated by combining the leaves of the tracks trees, some are selected to update tracks, and then the set of hypotheses is cleared. The advantage of the track-oriented approach is the easyness offered for maintaining and updating tracks compared to the difficulty of managing the set of global hypotheses over time. For instance it is shown in [19] that maintaining several hundreds of tracks and expanding them in hypotheses is feasible. In [20] a relatively simple breadth-first approach to hypothesis formation is proposed : it proceeds by iteratively extending hypotheses with potential tracks and stopping the process when the resulting score is too low compared to the best incomplete hypothesis formed until then.

One approach to make track-oriented MHT even more efficient could be to adapt Multi Dimensional Assignment (MDA) techniques [50, 133] to the multiple frame case. This idea has been referred several times [18] as a promising future development of the MHT. Adapting MDA to the multiple frame problem is however a delicate task since the multiple frame problem cannot be formulated in a straightforward way as a MDA problem. In contrast with the multidimensional association problem the costs of the associations are indeed not independent on each other in the multiple frame case. Instead the cost of linking two measurements depends on the whole association history of the tracks due to the recursive state estimation procedure. This theoretical issue may explain why to the best of our knowledge MDA has not been used yet in any track-oriented MHT algorithm. This is why we have opted for developing a new MHT implementation, relying on the track-oriented approach proposed in [20], but which takes benefits of the topology of the track trees to solve efficiently the multi frame data association problem, and including target perceivability concepts.

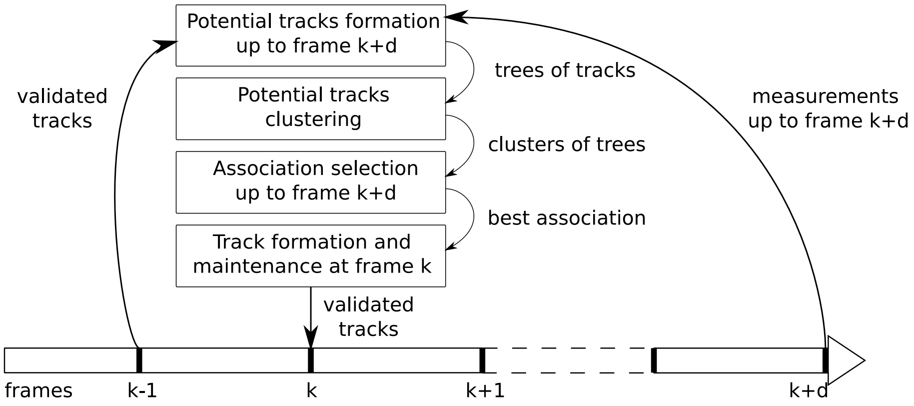


FIG. 7.1: The proposed E-MHT flow chart.

7.2 Enhanced MHT for particle tracking in fluorescence images processing

Despite numerous improvements the MHT approach has never been used for tracking biological targets tracking in fluorescence microscopy images because its computational demand has been deemed prohibitive. We introduce here an approach that should change this view since we propose a MHT algorithm that is efficient even when tracking numerous particles in dense clutter thanks to our algorithmic design and the use of advanced computing technologies. We call this algorithm E-MHT for *Enhanced* MHT.

7.2.1 Introducing target perceivability in the MHT

Frequent target appearance and disappearance events combined with the very high density of clutter can lead to continuation of tracks with spurious measurements and false tracks construction in applications such as bioimaging. Hence, it is essential to accurately model the target existence (see Chapter 6) and to include it into a multiple frame scheme such as the MHT to extract relevant information from the data. A proposition of a Probabilistic MHT (PMHT) [183] algorithm using a Markovian state of targets is introduced in [124], but some fundamental principles of the MHT are relaxed in order to reduce computational complexity. An example of this is that candidate tracks are averaged in place of the enumeration of the set of possible associations. In cluttered environments, the performance of basic PMHT techniques turns out to be at best similar to that of the joint probabilistic data association filter [13] algorithm which is much less numerically demanding. So in high clutter density and numerous targets conditions we argue that imposing a strict competition between tracks is a key advantage for discarding false measurements and detecting tracks initiation and termination events. Hence, in contrast with the PMHT approach, we want to impose that tracks compete, and we will therefore focus on including target perceivability in the original definition of the MHT algorithm.

7.2.2 A new fast MHT algorithm for particle tracking

The MHT technique we propose aims at building iteratively the association $\Theta^\star(k)$ that maximizes the likelihood $P\{\Theta^{k+d}|Z^{k+d}\}$, instead of maximizing $P\{\Theta^k|Z^k\}$ as instantaneous Bayesian algorithms do, whereby $d-1$ future frames are accounted for when selecting $\Theta^\star(k)$. We include the target perceivability concept presented in Chapter 6 in the estimation of tracks likelihood :

the probability $P\{\Theta^{k+d}|Z^{k+d}\}$ is computed by combining Equation 5.8 and Equation 6.5, while the sequential target perceivability estimation procedure described in Chapter 6 is applied over time.

We adopt a sequential approach which consists mainly in four steps for each frame of the image sequence. This procedure is summarized in Fig. 7.1. In practice, at frame k the previous processing steps provide the set of tracks $\Theta^{\star k-1}$ and we have to extend the tracks during $d+1$ frames with measurements from the set $Z^{k:k+d}$. To do so we first build independently for each track $\theta_j^{\star k-1} \in \Theta^{\star k-1}$ the set of potential tracks Γ_j^{k+d} up to frame $k+d$ formed by the possible associations of $\theta_j^{\star k-1}$ with detections from $Z^{k:k+d}$. The second step consists in dividing the global association problem into a set of smaller tasks by clustering the association trees that are concurrent for at least one measurement. We denote $\Gamma_{c_i}^{k+d}$ the set of potential tracks forming the cluster c_i. Then in the association selection step we build a subset of potential tracks $\Theta_{c_i}^{\star k+d} \subset \Gamma_{c_i}^{k+d}$ which has the highest likelihood $\mathcal{L}(\Theta_{c_i}^{k+d})$ for each cluster c_i. In this stage we ensure that the built solution complies with the unicity of association principle : each non-terminated track is associated to a single measurement, while a measurement is assigned to one track at most. In the final step the best association is built by merging the association found for each cluster : $\Theta^{\star k+d} = \cup_i \Theta_{c_i}^{\star k+d}$. On this basis, validated tracks are either continued or ended while a number of new tracks are validated. We detail next each of these four steps.

7.3 Potential tracks formation

At frame k the previous frames processing provides the set of tracks $\Theta^{\star k-1}$ which is a suboptimal solution of the problem consisting in maximizing $P\{\Theta^{k-1}|Z^{k-1}\}$. The next step of the MHT algorithm is to find a set of tracks $\Theta^{\star k+d}$ maximizing $P\{\Theta^{k+d}|Z^{k+d}\}$, the solution $\Theta^{\star k+d}$ being formed by extending $\Theta^{\star k-1}$ and creating new tracks with detections from the set of measurements $Z^{k:k+d}$. To do so, we first build independently the potential associations of each track in $\Theta^{\star k-1}$ with measurements from $Z^{k:k+d}$, and we also create one track for each target appearance event.

7.3.1 Continuing existing tracks

At frame k each track $\theta_j^{k-1} \in \Theta^{\star k-1}$ can be associated either to a detection in $Z(k)$ or to a virtual detection modeling its temporary disappearance. The association with detections in $Z(k)$ is restricted to a subset of measurements falling into the track search gate, which reduces considerably the complexity of the track enumeration procedure. The potential associations at frame k form a set $\Gamma_j^k = \{\theta_{j,z_i}^k\}_{i=1..p}$ of tracks where we denote θ_{j,z_i}^k the potential track built by associating the track θ_j^{k-1} with the detection z_i which is either virtual or belonging to $Z(k)$.

We iteratively repeat the association process for every set Γ_j^t with $t = k..k+d-1$, hence the formation of potential tracks Γ_j^{k+d} can be modeled as the construction of trees of feasible associations. Indeed, a track θ_j^{k-1} gives birth to a set of potential tracks Γ_j^k which in turn create potential tracks and so on. The creation of a tree of associations to represent the potential tracks originating from a track $\theta_j^{\star k-1}$ is illustrated with a simple example showing few detections in two subsequent frames in Figure 7.2. We show that a single measurement can be represented by multiple nodes since it can be associated to multiple potential tracks. Finding the same detection in different nodes of set of trees creates numerous incompatibilities between branches of the association trees since tracks cannot share detections.

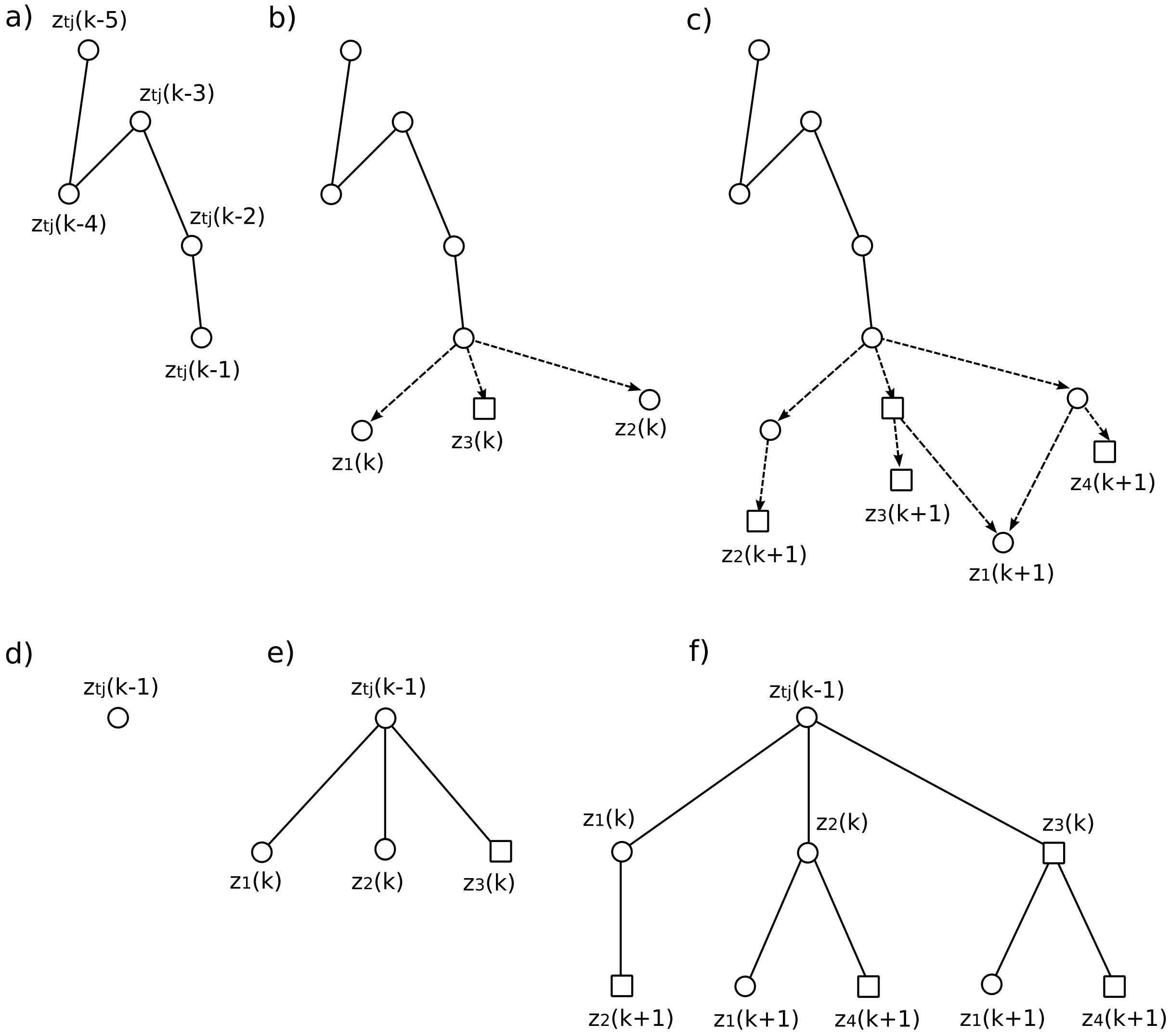


Fig. 7.2: Construction of a tree of potential tracks in two subsequent frames. Real measurements are represented by a circle and virtual measurements by a square. a) a validated track $\theta_j^{\star k-1}$. b) the construction of Γ_j^k which is composed of three potential tracks to continue $\theta_j^{\star k-1}$ at time k : two feasible associations with real measurements $z_1(k)$ and $z_2(k)$ and and one with the virtual measurement $z_3(k)$. c) extensions of Γ_j^k to time $k+1$: potential tracks associate with the only real measurements $z_1(k+1)$ and with a predicted measurements which is specific to each potential track. The track $\theta_{jz_1(k)}^{\star k-1}$ does not associate with $z_1(k+1)$ because the measurement does not fall inside the search gate of the track. d), e) and f) tree representation of the associations displayed in figures a), b) and c), respectively. The leaves correspond to the set of potential tracks Γ_j^{k+1} which emerges from $\theta_j^{\star k-1}$.

7.3.2 Modeling the appearance of new targets

We create one tree of associations $\Gamma_j^{k:k+d}$ for each non terminated track $\theta_j^{k-1} \in \Theta^{\star k-1}$, whose root node is θ_j^{k-1}, in order to model the mandatory continuation of tracks from $\Theta^{\star k-1}$. Moreover, one tree is created from each detection in $Z^{k:k+d}$ to account for the possibility for new targets to appear. This last feature considerably increases the number of trees : we show in Figure 7.3 that 6 trees are required to model a fairly simple tracking problem for two frames : two non-terminated tracks already exist at frame $k-1$ and two detections are observed in each frame k and $k+1$. The appearance of new targets is represented in this case by 4 trees over the 6 built trees. We denote in the following $\Gamma^{k:k+d}$ the set of potential tracks corresponding to either the continuation of an existing track of to the appearance or a new target.
While increasing the complexity of the association selection procedure, systematically creating a new tree for each detection constitutes a fully automatic track creation process without any *ad hoc* decision procedure : the appearance of new targets events are detected by the association selection procedure. Indeed, selecting at time k one leaf from the tree originating from the detection z_i results in the creation of a new track which aims at following the putative target from which z_i originates (see trees maintenance step described in Section 7.5). On the contrary, the track creation is rejected and z_i is assumed to originate either from an already existing target or from clutter if no leaf is selected in this tree. The main benefit of the approach is the guarantee of the optimality at frame k of the built solution $\Theta^{\star k+d}$ regarding the score function $\mathcal{L}(\Theta^{\star k+d})$. Target appearance is indeed decided only when the track creation yields an increase of $\mathcal{L}(\Theta^{\star})$. By doing so the decision process accounts for the target appearance statistics, as well as the statistical properties of the false detections and the stochastic behavior of the existing targets.

7.3.3 Perceivability based track labeling

During the nodes formation process we label the potential tracks according to their probability of perceivability as introduced in Chapter 6 : a track θ_j^t at frame t is *confirmed* if $\exists t' \leq t$ such that $\lambda_j^1(t'|t') \geq p_c$, and *terminated* if $\lambda_j^1(t+1|t) \leq p_t$. Extending a track with a low probability of perceivability is useless since it is very unlikely that it will produce some detections in the following frames. We thus stop to associate more measurements to a terminated track. As an example, in Figure 7.3 the potential track associated to the virtual detection $z_3(k)$ is stopped, which results in a simplified association problem since no more potential tracks originate from this node. Track confirmation plays also an important role to reduce the cost of the association hypothesis selection step as we will show in Section 7.5.
Nodes in $\Gamma_j^{k:k+d}$ without any link to a node at the next level of the tree are coined *leaves* and constitute the possible tracks for continuing the track θ_j^{k-1}. One leaf corresponds either to a terminated track, or to a track with associations up to time $k+d$, hence reaching the bottom of the tree.

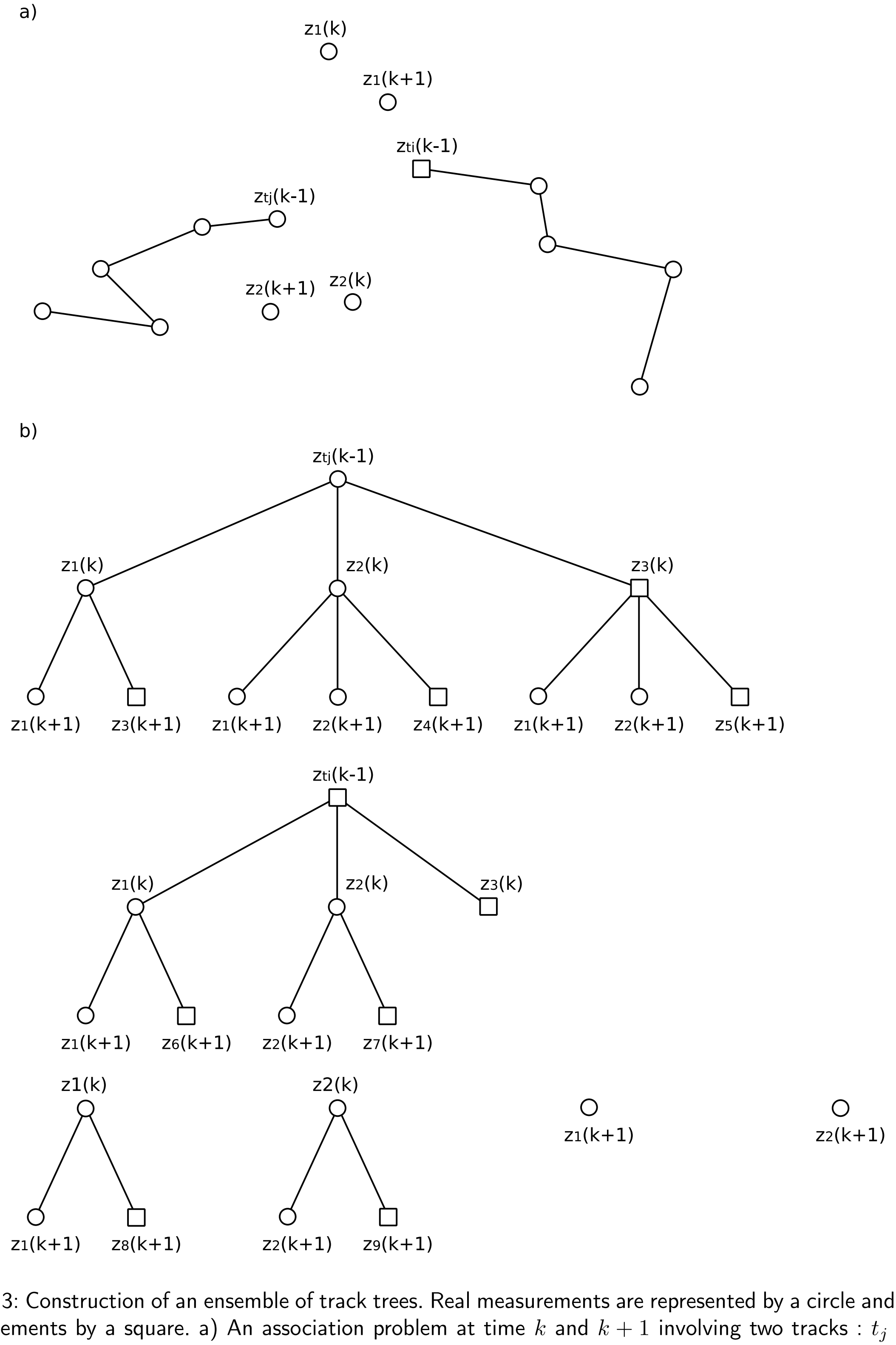


FIG. 7.3: Construction of an ensemble of track trees. Real measurements are represented by a circle and virtual measurements by a square. a) An association problem at time k and $k+1$ involving two tracks : t_j and t_i. b) the set of track trees corresponding to the association problem a). One tree is built for each track t_i and t_j and four trees are built for each possible target appearance events. It results in trees with root nodes at different times : $k-1$, k and $k+1$. The overall complexity is reduced by measurements gating and also by track termination. For instance the track $\theta^{k-1}_{t_i z_3(k)}$ is not extended with any detection at time k because of its low predicted probability of perceivability at time $k+1$, hence resulting in a leaf at time k.

7.3.4 Assembly of the method

At each frame k we aim at building a set of trees for modeling existing tracks continuation and new targets appearance. However, this process can be computationally intensive if the number of measurements is large. In order to reduce the cost of the potential tracks construction we re-use the trees built during the previous frame processing since most of the nodes are common to subsequent iterations and only few changes have to be made :

1. existing leaves at time $k+d-1$ give birth to a set of leaves at time $k+d$ when extending them with the measurements from the set $Z(k+d)$,
2. new trees must be created from measurements in $Z(k+d)$ to model target appearance events in the frame $k+d$,
3. trees corresponding to target appearance events in frame $k-1$ must be discarded if they are not selected by $\Theta^{\star k-1}$. The tree cleaning process is achieved in practice during the track update stage (see Section 7.5.8).

In practice the procedure is very efficient : we demonstrate in Section 7.6 that the computational demand of this step is nearly negligible compared to the hypothesis selection stage.

7.4 Potential tracks clustering

The goal of clustering potential tracks is to reduce the complexity of the data association problem by dividing it into a set of smaller problems that can be solved independently from each other. A set of tracks Γ^{k+d} can be clustered if each subset of tracks $\Gamma^{k+d}_{c_i} \subset \Gamma^{k+d}$ do not use any measurement associated to tracks from other clusters. In practice tracks from the same tree are all concurrent since they share at least one measurement : the root node. So we cluster trees instead of tracks, which significantly reduces the complexity of the clustering task since trees are far less numerous than leaves. To do so we use a bottom-up approach to clustering : first each tree is considered as a cluster, and then clusters are merged until no more shared measurements remain.

Clustering the association problem is a very efficient technique to reduce the data association cost. Indeed, the cost of solving two independent problems with similar complexity is twice the cost of solving one such problem. In contrast, solving twice a bigger problem is far more computationally intensive than just double since the computational and memory demand of enumerative solving techniques grows exponentially with the number of measurements. Moreover, independent association problems can be solved in parallel, hence, the time spent to solve the two problems can be as low as the processing time of the most complex one.[4].

7.5 Creating the optimal association hypothesis up to frame $k+d$

7.5.1 Principles of the novel implementation

The association hypothesis selection procedure consists in finding a subset of potential tracks $\Theta^{\star k+d}$ maximizing the likelihood $\mathcal{L}(\Theta^{k+d})$. Enumerating all the tracking solutions is however unfeasible in a short time because of their larger number which is significantly increased by the presence of false detections that should remain not associated to any track. Fortunately the unicity principle of association between tracks and detections imposes numerous constraints that

---

[4] This is of course a simplistic view which does not account for the number of available parallel computing units and other parallel tasks

can be exploited to reduce the number of considered associations. A usual way to solve the issue is to formulate it as a standard convex optimization problem with a set of constraints given by track incompatibilities [20]. We have however preferred to develop our own solver because it can be made to fully take advantage of the tree structure of the tracks formation process in two ways :

1. by building only a very restricted number of associations Θ^{k+d} by pruning huge sets of solutions based on the tree structure of the potential tracks,
2. by exploring and combining the trees in a massively parallel way.

The benefits of the proposed implementation over the breadth-first implementation introduced in [20] are twofolds : a greater efficiency by exploiting the topology of the solutions, and the recovery of the exact solution $\Theta^{\star k+d}$ since the pruning technique we use is exact instead of being heuristic.

7.5.2 Hypothesis score

Track-oriented score function

At frame k we re-write the likelihood of one feasible association Θ^{k+d} as :

$$\mathcal{L}(\Theta^{k+d}) \propto \prod_{t=1..k+d} p\{Z_0(t)\} \prod_{\theta_j \in \Theta^{k+d}} p(\theta_j^{k+d}, z_{t_j}^{k+d}), \tag{7.1}$$

by exploiting the independence property of tracks (Eq. 5.8). As indicated by Equation 6.5 the joint probability $p(\theta_j^{k+d}, z_{t_j}^{k+d}) = \chi_j^{k+d}$ for each track $\theta_j \in \Theta^{k+d}$ can be estimated in a recursive way. We thus split the computation of this probability in two terms : the joint probability of the track up to time $t-1$, and the conditional probability up to time $t+d$:

$$\chi_j^{k+d} = \chi_j^{k-1} p(\theta_j^{k:k+d}, z_{t_j}^{k:k+d} | \theta_j^{k-1}, z_{t_j}^{k-1}). \tag{7.2}$$

We denote next $\chi_j^{k:k+d} \triangleq p(\theta_j^{k:k+d}, z_{t_j}^{k:k+d} | \theta_j^{k-1}, z_{t_j}^{k-1})$ for the second term. Using these notations we write Equation 7.1 as :

$$\begin{align} \mathcal{L}(\Theta^{k+d}) &\propto \prod_{t=1..k+d} p\{Z_0(t)\} \prod_{\theta_j \in \Theta^{k+d}} \chi_j^{k-1} \chi_j^{k:k+d}, \tag{7.3} \\ &\propto \mathcal{L}(\Theta^{\star, k-1}) p\{Z_0^{k:k+d}\} \prod_{\theta_j \in \Theta^{k+d}} \chi_j^{k:k+d}, \tag{7.4} \end{align}$$

Each term $\chi_j^{k:k+d}$ corresponds to the conditional probability of the leaf θ_j^{k+d} chosen in the track tree $\Gamma_j^{k:k+d}$, while the set $Z_0^{k:k+d}$ contains the measurements that are not used by the selected leaves in frames k to $k+d$. When solving the tracking problem at frame k the set of tracks $\Theta^{\star, k+d}$ is already fixed, so, building the optimal association $\Theta^{\star, k+d}$ consists in selecting a subset of leaves in $\Gamma^{k:k+d}$ maximizing the following score function :

$$\mathcal{F}(\Theta^{k+d}) = p\{Z_0^{k:k+d}\} \prod_{\theta_j \in \Theta^{k+d}} \chi_j^{k:k+d}. \tag{7.5}$$

A normalized hypothesis score for efficient hypothesis pruning

When the detection $z(t)$ is associated to one track θ_j, as shown in Chapter 6, the joint probability χ_j^t is computed as the product of the previous joint probability χ_j^{t-1} by an update factor which is dependent on $z(t)$:

$$\chi_j^t = \chi_j^{t-1}\left[\lambda_j^0(t|t-1)p_{FD}(z_{t_j}(t),t) + \lambda_j^1(t|t-1)P_D P_G g(z_{t_j}(t)|\theta_j^t, z_{t_j}^{t-1})g(\theta(t)|\theta_j^{t-1}, z_{t_j}^{t-1})\right]. \quad (7.6)$$

We denote next $\Xi_j(z_{t_j}(t))$ the update factor such that $\chi_j^t = \chi_j^{t-1}\Xi_j(z(t))$ if the detection $z_{t_j}(t)$ is associated to the track θ_j. With these notations the conditional joint probability up to time $k+d$ can be re-written as the product of $k+1$ factors :

$$\chi_j^{k:k+d} = \prod_{t=k..k+d} \Xi_j(z_{t_j}(t)). \quad (7.7)$$

On the other hand, the tracks score is multiplied by $p_{FD}(z(t),t)$ if the detection $z(t)$ is not associated to any track because $z(t)$ is considered as being a false detection (Eq. 5.31) in this case. By combining Equation 7.7 with the false detection probability we thus re-write the score function $\mathcal{F}(\Theta^{k+d})$ as follows :

$$\mathcal{F}(\Theta^{k+d}) = \prod_{z(t)\in Z_0^{k:k+d}} p_{FD}(z(t),t) \prod_{\theta_j\in\Theta^{k+d}} \prod_{t=k..k+d} \Xi_j(z_{t_j}(t)). \quad (7.8)$$

For a feasible solution Θ^{k+d} every detection $z(t) \in Z^{k:k+d}$ is either used by a tracking hypothesis or set to the false detection set $Z_0^{k:k+d}$. Every measurement $z(t) \in Z^{k:k+d}$ therefore participates to the value of $\mathcal{L}(\Theta^{k+d})$ by means of a multiplicative factor, as shown by Equation 7.8. In the following we denote $\beta(z(t))$ the largest factor which can be induced by the presence of a detection $z(t)$ in the data set :

$$\beta(z(t)) \triangleq \max\left(p_{FD}(z(t),t), \{\Xi_j(z(t))\}_{\theta_j}\right), \quad (7.9)$$

where we consider all tracks θ_j associated to the measurement $z(t)$. When computing $\mathcal{L}(\Theta^{k+d})$ the tracks joint probability χ^{k-1} is therefore multiplied for each detection $z(t) \in Z^{k:k+d}$ by a factor which is at most $\beta(z(t))$. We therefore deduce that the leaves score in frame $k+d$ cannot exceed the following upper bound :

$$\mathcal{F}(\Theta^{k+d}) \leq \prod_{z(t)\in Z^{k+d}} \beta(z(t)). \quad (7.10)$$

It is worth noting that the upper bound $\prod_{z(t)\in Z^{k+d}} \beta(z(t))$ does not in general correspond to a feasible set of leaves. As an example, assuming that the maximum factor $\beta(z(t))$ for a detection $z(t) \in Z^{k:k+d}$ corresponds to the score $\Xi_j(z(t))$ of associating the detection with a potential track θ_j, other associations in θ_j are not guaranteed to be selected by other maximum factors.
The product $\prod_{z(t)\in Z^{k+d}} \beta(z(t))$ being constant across the feasible set of associations we define a normalized score function $\hat{\mathcal{F}}(\Theta^{k+d})$:

$$\hat{\mathcal{F}}(\Theta^{k+d}) \triangleq \mathcal{F}(\Theta^{k+d})\left(\prod_{z(t)\in Z^{k+d}} \beta(z(t))\right)^{-1} \quad (7.11)$$

which is proportional to $\mathcal{L}(\Theta^{k+d})$ and inferior to 1 (see Eq. 7.10). We develop the normalized score as :

$$\hat{\mathcal{F}}(\Theta^{k+d}) = \prod_{z(t)\in Z_0^{k:k+d}} \frac{p_{FD}(z(t),t)}{\beta(z(t))} \prod_{\theta_j\in\Theta^{k+d}} \prod_{t=k..k+d} \frac{\Xi_j(z_{t_j}(t))}{\beta(z_{t_j}(t))}, \tag{7.12}$$

$$= \prod_{z(t)\in Z_0^{k:k+d}} \hat{p}_{FD}(z(t),t) \prod_{\theta_j\in\Theta^{k+d}} \prod_{t=k..k+d} \hat{\Xi}_j(z_{t_j}(t)), \tag{7.13}$$

which highlights that $\hat{\mathcal{F}}(\Theta^{k+d})$ can be computed as the product of one normalized term for each measurement. The factor associated to a real detection $z(k)\in Z^{k+d}$ is either $\hat{p}_{FD}(z(t),t)$ if the measurement is not associated to any track, or $\hat{\Xi}_j(z_{t_j}(t))$ if it is associated to the track θ_j. In every case this factor is inferior to 1 by definition of $\beta(z(t))$ (Eq. 7.9). The same property holds for a virtual measurement $z(k)$ which is created to extend a track θ_j despite the lack of one measurement : if the virtual measurement is not associated to θ_j we give it the probability 1, and $\Xi_j(z(t))$ otherwise. Hence, for such a virtual detection $\beta(z)=1$ and we can omit its value when normalizing the tracking score function.

The normalization of the probability values can be expressed at a single track level by defining a normalized conditional joint probability for each leaf θ_j as :

$$\hat{\chi}_j^{k:k+d} \triangleq \prod_{t=k..k+d} \hat{\Xi}_j(z_{t_j}(t)), \tag{7.14}$$

which closely resembles the non-normalized one (Eq. 7.7). Using these notations the normalized score function $\hat{\mathcal{F}}(\Theta^{k+d})$ is re-written as :

$$\hat{\mathcal{F}}(\Theta^{k+d}) = \prod_{z(t)\in Z_0^{k:k+d}} \hat{p}_{FD}(z(t),t) \prod_{\theta_j\in\Theta^{k+d}} \hat{\chi}_j^{k:k+d}, \tag{7.15}$$

which also has a very similar structure to the non-normalized score function $\mathcal{F}(\Theta^{k+d})$ as defined in Equation 7.5. The similar structure of both scores allows us to define next an optimization procedure which is suited to both scores, without regard for the possible normalization of the probabilities. We will however show that using the normalized score yields a more efficient search procedure because of the stricter hypotheses pruning criterion which is implied in this case.

7.5.3 Algorithm initialization with a local solution

Frame-to-frame tracking algorithms accounting for only one frame at a time do not provide in general the optimal solution of the tracking problem as formulated in the MHT framework (Eq. 7.1). Such algorithms have however proved to give good quality results in a number of biological studies [6, 94]. We therefore propose to apply such a simplified approach between frames k and $k+d$ in order to build a solution $\tilde{\Theta}^{k+d}$ which closely resembles the hypothesis $\Theta^{\star k+d}$. This first approximated solution will allow the optimal tracks search procedure to discard rapidly huge sets of solutions which have significantly lower score.

The simplified procedure we propose consists in the following stages :

1. At frame k we browse through the sets of nodes linked to each validated track $\theta_j^{\star k-1}\in\Theta^{\star k-1}$, and we select θ_j^k, the node with the largest score $\Xi_j(z(k))$.
2. We repeat the search procedure (1) with the remaining detections and nodes until every node in $\Theta^{\star k-1}$ is associated to a measurement.

3. At this stage some detections may remain unassociated. We simplify here the tracking problem by assuming that no new target appears and we consider each non associated detection as being a false detection. This default choice is supported by the probability of false detection P_{FD} which is in general superior to the probability of new target P_{NT}.
4. The selected nodes are collected in the set $\tilde{\Theta}^k$.

This procedure is then iterated through time, extending $\tilde{\Theta}^t$ at each time t (instead of $\Theta^{\star k-1}$) with nodes from Γ^{t+1}, until a feasible solution $\tilde{\Theta}^{k+d}$ is formed.

The proposed scheme to obtain the tracking hypothesis $\tilde{\Theta}^{k+d}$ is remarkably non expensive, indeed, all the time consuming steps : associations construction, memory allocation, state space estimation computations, score computations, etc, have been precomputed for the need of the MHT algorithm. We will however show with a number of experiments in Section 7.6 that the procedure provides a remarkably good approximation of the optimal solution $\Theta^{\star k+d}$ despite its low cost. It indeed benefits from the statistical framework for computing scores as well as from the robust track termination/confirmation process.

At this stage of the optimization scheme $\tilde{\Theta}^{k+d}$ is the only feasible solution at hand and we denote $\hat{\mathcal{F}}(\tilde{\Theta}^{k+d})$ its normalized score.

7.5.4 Track score back-propagation

Before solving the tracking problem we first back-propagate the leaf normalized scores to upper nodes such that each node $\theta_j \in \Gamma^{k:k+d}$ contains the information of the score $\hat{\chi}^{\star}(\theta_j)$ which is the largest among the leaves of the tracks subtree originating from it. The recursive procedure for back-propagating scores is illustrated in Figure 7.4 with a simple example.

Thanks to the back-propagation step we are able to know beforehand the greatest score we would be able to achieve by going through a node when exploring a tracks tree. This feature is the basis of an efficient subtree pruning technique described next.

7.5.5 Combining trees of association

The optimization procedure starts by selecting an arbitrary track θ_1 in the first tree. Hence, for now the set of tracks is $\Theta = \{\theta_1\}$ and we compute the incomplete score of this hypothesis as : $\tilde{\mathcal{F}}(\Theta) = \hat{\chi}_1^{k:k+d}$. The second tree Γ_2^{k+d} is then considered and two pruning techniques are used to check its compatibility with the current hypothesis Θ.

The compatibility of the root node of Γ_2^{k+d}, say θ_2, is first checked against the detections used by Θ :

- if the measurement associated to θ_2 is already used by any track in Θ the hypothesis is abandoned since a measurement can be associated to one track at most,
- otherwise, the tree originating from θ_2 is validated for the second pruning step.

If the node θ_2 is rejected by this first compatibility test, then the whole subtree of tracks originating from the node is rejected (Γ_2^{k+d} in this case). The next tree, for instance Γ_3^{k+d}, is considered for association in this unfavorable case. We consider instead here that the node θ_2 is validated by the criterion of detections compatibility.

The largest score among leaves originating from the current node θ_2 is directly available at this stage thanks to the prior scores backpropagation : the best achievable score in the whole subtree is $\chi^{\star}(\theta_2)$. We denote $\tilde{\mathcal{F}}^{\star}(\Theta, \theta_2)$, the highest partial score when adding one leaf from the tree originating from θ_2 to the hypothesis Θ, and we write :

$$\tilde{\mathcal{F}}^{\star}(\Theta, \theta_2) \triangleq \tilde{\mathcal{F}}^{\star}(\Theta)\hat{\chi}^{\star}(\theta_2). \tag{7.16}$$

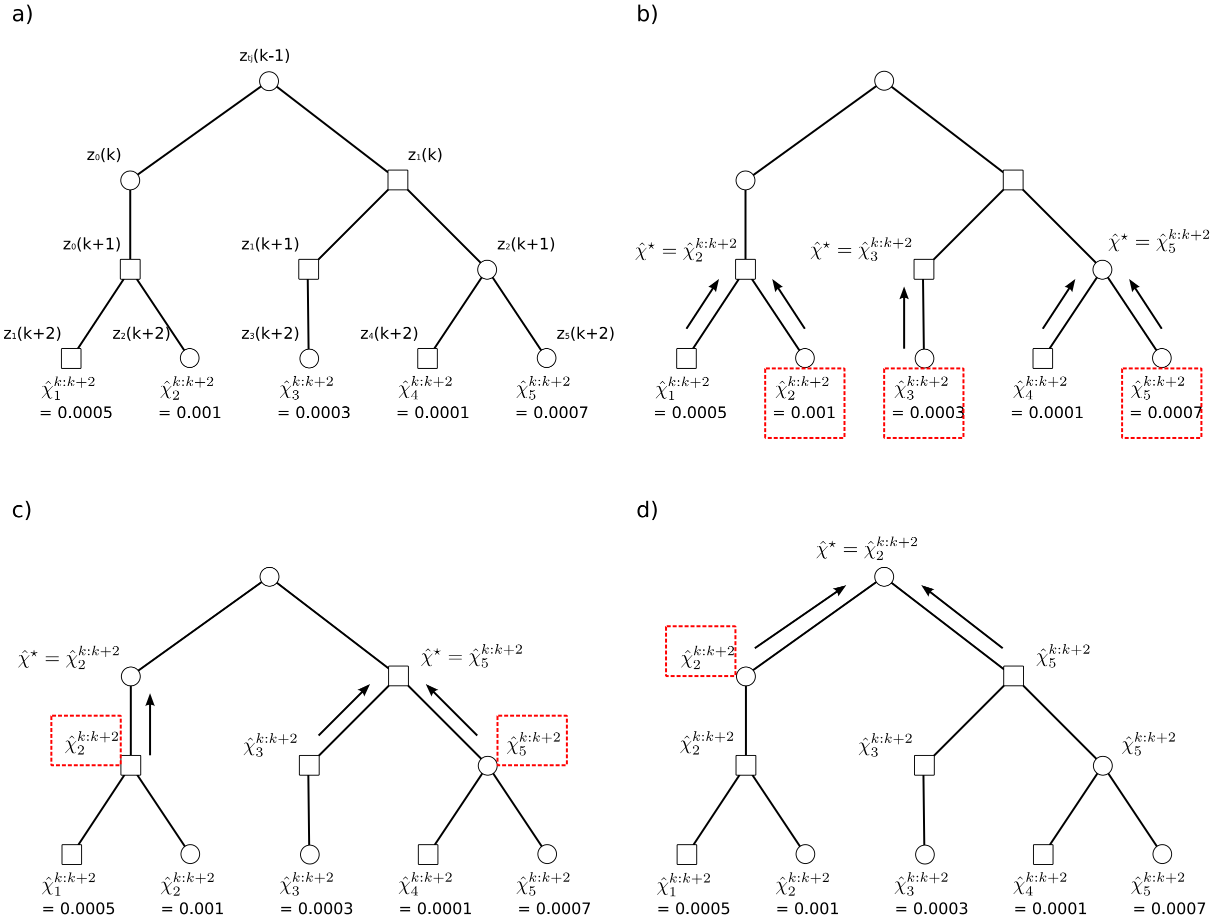


FIG. 7.4: Normalized track scores back-propagation in a tracks tree. a) The tracks tree showing the associations and the normalized score of the leaves. Real measurements are represented by a circle and virtual measurements by a square. b) Back-propagation of the best scores at time $k+2$ (circled in red) to the nodes at time $k+1$. c) Back-propagation of the best scores at time $k+1$ to nodes at time k. d) Back-propagation of the best score of the tree to the root node. Every node in the tree contains the information of the best leaf score in the subtree originating from it.

We assume here that the leaf θ_j, corresponding to the maximum score $\chi^\star(\theta_2)$, is compatible with the detections already used by Θ. However, combining θ_j with Θ does not necessarily provides a feasible solution since some tracks $\theta_j^{k-1} \in \Theta^{\star k-1}$ may still have to be continued. Moreover, the probability of the set of non assigned detections is not accounted for in the incomplete score $\tilde{\mathcal{F}}^\star(\Theta, \hat{\theta}_2^{k-1})$. Accounting for additional tracks and false detections to make $\Theta \cup \theta_j$ a feasible solution results in a score $\hat{\mathcal{F}}(\Theta)$ which is obtained by multiplying the partial score $\tilde{\mathcal{F}}^\star(\Theta, \theta_2)$ by the score of the missing tracks and the score of the false detections, as shown in Equation 7.15 (or Eq. 7.5 for non-normalized scores). These missing multiplicative factors are all below 1, we thus infer that the partial score $\tilde{\mathcal{F}}^\star(\Theta, \theta_2)$ is an upper bound of the score of the tracking solutions which can be generated by using Θ and θ_2. We can therefore write :

$$\hat{\mathcal{F}}(\Theta') \leq \tilde{\mathcal{F}}^\star(\Theta, \theta_2), \tag{7.17}$$

for every tracking hypothesis Θ' containing both Θ and θ_2. When considering the association with the node θ_2, if the upper bound $\tilde{\mathcal{F}}^\star(\Theta, \theta_2)$ falls beyond the score of the best hypothesis built until now $\tilde{\Theta}^{k+d}$ we deduce by transitivity of the score inequality that the score of every tracking solution Θ' containing Θ and θ_2 is below $\hat{\mathcal{F}}(\tilde{\Theta}^{k+d})$:

$$\hat{\mathcal{F}}(\Theta') \leq \tilde{\mathcal{F}}^{\star}(\Theta, \theta_2) \leq \hat{\mathcal{F}}(\tilde{\Theta}^{k+d}). \tag{7.18}$$

In such a case continuing the search procedure with nodes linked to the current node θ_2 is useless since it would result in solutions with lower scores than $\tilde{\Theta}^{k+d}$. We thus reject the extension of Θ with θ_2 and choose another feasible node in place of θ_2 to extend Θ.

If the current node θ_2 satisfies both the compatibility test with Θ and the bounding test on $\tilde{\mathcal{F}}^{\star}(\Theta, \theta_2)$ we repeat the exploration procedure for every node it is linked to at the next level. This process is applied recursively until the end of a branch is reached. In this case a Θ duplicate is extended with the track θ_j with which we ended : $\Theta = \Theta \cup \theta_j$, and another tree is selected.

An hypothesis Θ is feasible if it contains a potential track to continue every non terminated track in $\Theta^{\star k-1}$. In this case, if the score $\hat{\mathcal{F}}(\Theta)$ is greater than $\hat{\mathcal{F}}(\tilde{\Theta})$, then Θ is stored as the best hypothesis until then : $\tilde{\Theta} = \Theta$. While being feasible, the hypothesis Θ is eventually extended with other potential tracks corresponding to appearance of targets. Indeed, using the non associated detections to form new tracks is a mean of substantially improving the hypothesis score $\hat{\mathcal{F}}(\Theta)$.

The exploration procedure is stopped as soon as no more feasible node remains. The two techniques we use for hypothesis pruning are exact, instead of heuristic, so we end up with certainty with the optimal solution : $\Theta^{\star k+d} = \tilde{\Theta}^{k+d}$. By taking advantage of the tree structure of the potential tracks we exclude huge sets of solutions each time a node is branched to the current hypothesis, which significantly speeds up the construction of the feasible associations Θ^{k+d}, but we still ensure the exact recovery of the optimal solution.

7.5.6 Parallel implementation

The proposed algorithm design for solving the MHT problem is fully compatible with massive parallel computing. Indeed, exploring branches of trees can be done in an independent manner. Hence, in practice we launch a new computing thread for each of these tasks. Parallel computing with a single computer having multiple computing units however requires an efficient management of the number of running computing tasks in order to optimize the use of the computing units and avoid the explosion of the memory requirements. To do so, we manage the thousands of threads with an optimized waiting queue which ensures that only few of them are running simultaneously, according to the number of available computing units. We show in Section 7.6 that the computation time of the proposed MHT implementation decreases almost linearly with the increase of the number of computing units, which reflects the optimality of the parallel design. Because it fully exploits parallel computing with a single computer, the proposed MHT implementation will take full advantage of the next generations of CPUs for which the number of cores is planned to increase considerably.

7.5.7 Why normalizing scores improves the search procedure efficiency

We have shown above that $\tilde{\mathcal{F}}(\Theta, \theta_i)$ is an upper bound on the achievable normalized scores when extending the tracking hypothesis Θ with the node θ_i. With the same reasoning it can be proved that a similar result holds when using the track scores instead of the normalized scores. Indeed, the highest partial score $\tilde{\mathcal{F}}^{\star}_{nn}(\Theta, \theta_i)$ computed with non normalized scores is still an upper bound of the score of feasible solutions including Θ and θ_i since computing the score of a complete solution involves multiplying $\tilde{\mathcal{F}}^{\star}_{nn}(\Theta, \theta_i)$ by additional factors which are all above 1.

Without loss of generality we consider Θ' which is a complete tracking solution including the tracking hypotheses in Θ and the best potential track going through the node θ_j. By combining Equation 7.5 and the definition of $\tilde{\mathcal{F}}^{\star}_{nn}(\Theta, \theta_i)$ we re-write the score of the hypothesis as :

$$\mathcal{F}(\Theta') = \tilde{\mathcal{F}}^{\star}_{nn}(\Theta, \theta_i) p\{Z_0^{k+d}\} \prod_{\theta \notin \Theta \cup \theta_i} \chi_i^{k:k+d}. \tag{7.19}$$

The multiplicative term $p\{Z_0^{k+d}\} \prod_{\theta \notin \Theta \cup \theta_i} \chi_i^{k:k+d}$ is below 1, which confirms that the non normalized partial score $\tilde{\mathcal{F}}^{\star}_{nn}(\Theta, \theta_i)$ is an upper bound of $\mathcal{F}(\Theta')$. Using Equations 7.7 and 7.8 we further decompose $\mathcal{F}(\Theta')$ as :

$$\mathcal{F}(\Theta') = \tilde{\mathcal{F}}^{\star}_{nn}(\Theta, \theta_i) \prod_{z(t) \in Z_0^{k+d}} p_{FD}(z(t), t) \prod_{\theta \notin \Theta \cup \theta_i} \prod_{t=k..k+d} \Xi_j(z_{t_j}(t)). \tag{7.20}$$

The most right-hand side of the equation is the product of terms corresponding to detections which are not accounted for by the incomplete tracking hypothesis $\Theta' \cup \theta_i$. For one such detection $z(t)$ the multiplicative term is by definition bounded by $\beta(z(y))$ (Eq. 7.9) which is also below 1. Hence the following inequalities hold :

$$\mathcal{F}(\Theta') \leq \tilde{\mathcal{F}}^{\star}_{nn}(\Theta, \theta_i) \prod_{z(t) \notin \tilde{Z}^{k:k+d}} \beta(z(t)) \leq \tilde{\mathcal{F}}^{\star}_{nn}(\Theta, \theta_i), \tag{7.21}$$

where we denote $\tilde{Z}^{k:k+d}$, the set of detections which is accounted for by the incomplete hypothesis $\Theta' \cup \theta_i$. The Equation 7.21 proves that the product $\tilde{\mathcal{F}}^{\star}_{nn}(\Theta, \theta_i) \prod_{z(t) \notin \tilde{Z}^{k:k+d}} \beta(z(t))$ is a tighter bound of $\mathcal{F}(\Theta')$ than $\tilde{\mathcal{F}}^{\star}_{nn}(\Theta, \theta_i)$. A criterion based on this quantity instead of $\tilde{\mathcal{F}}^{\star}_{nn}(\Theta, \theta_i)$ would thus result in a more efficient pruning techniques since we would be able to detect earlier hypotheses which best achievable scores fall beyond the threshold $\mathcal{F}(\tilde{\Theta})$. Computing this tight bound may be however time consuming since every single detection has to be tested against $\Theta \cup \theta_i$.

We re-write the tight bound of $\mathcal{F}(\Theta')$ as :

$$\tilde{\mathcal{F}}^{\star}_{nn}(\Theta, \theta_i) \prod_{z(t) \notin \tilde{Z}^{k:k+d}} \beta(z(t)) = \tilde{\mathcal{F}}^{\star}_{nn}(\Theta, \theta_i) \frac{\prod_{z(t) \in Z^{k:k+d}} \beta(z(t))}{\prod_{z(t) \in \tilde{Z}^{k:k+d}} \beta(z(t))}, \tag{7.22}$$

and since the best partial normalized score $\tilde{\mathcal{F}}^{\star}(\Theta, \theta_i)$ can be written as :

$$\tilde{\mathcal{F}}^{\star}(\Theta, \theta_i) = \frac{\tilde{\mathcal{F}}^{\star}_{nn}(\Theta, \theta_i)}{\prod_{z(t) \in \tilde{Z}^{k:k+d}} \beta(z(t))}, \tag{7.23}$$

from Equation 7.11, we then deduce a second expression of the tight bound on $\mathcal{F}(\Theta')$:

$$\mathcal{F}(\Theta') \leq \tilde{\mathcal{F}}^{\star}(\Theta, \theta_i) \prod_{z(t) \in Z^{k:k+d}} \beta(z(t)). \tag{7.24}$$

Dividing each side of the inequality by the constant $\prod_{z(t) \in Z^{k:k+d}} \beta(z(t))$ yields the following expression :

$$\hat{\mathcal{F}}(\Theta') \leq \tilde{\mathcal{F}}^{\star}(\Theta, \theta_i), \tag{7.25}$$

which shows that using the standard pruning criterion with the normalized track scores corresponds to using the tight bound for pruning with the non-normalized scores. Using the normalized

scores therefore results in stricter pruning decisions which discard more tracking hypotheses, whereby the efficiency of the optimal solution search procedure is improved.
We note that the same optimization trick cannot be applied to the normalized scores since an upper bound below 1 cannot be found for the normalized multiplicative factor corresponding to one detection : by definition there exists an association decision for which the normalized factor value is 1. Other ways have thus to be thought of to design tighter bounds of hypotheses scores which would improve even more the efficiency of the optimization algorithm.

7.5.8 Trees maintenance

Once the optimal association up to time $k + d$ has been built a number of maintenance operations have to be performed to the set of track trees. Indeed numerous potential tracks are not feasible anymore because of their incompatibility with the associations in $\Theta^\star(k)$. To do so we use a technique coined *N-scan pruning* in [20] whose principle is to identify root nodes at time $k - 1$ and cut all branches directly linked to them that do not correspond to $\Theta^\star(k)$ since they are not feasible. This process results in an important simplification of the track trees : only one branch is maintained for a tree with many branches linked to the root node since only one branch can use the common node at time k. Subtrees originating from other branches are therefore cleared. As an example we show in Figure 7.5 the maintenance of the set of trees representing the tracking problem presented in Figure 7.3. We show that updating the trees according to the optimal solution $\Theta^{\star k+d}$ results in keeping only 9 leaves over the original 18 potential tracks in this case. The maintenance operation thus avoids the exponential explosion of the number of leaves over time.
Detecting target appearance/disappearance events is also achieved during the trees maintenance stage. For instance, a track selected by $\Theta^{\star k+d}$, and with a root node found at time k, corresponds to a target appearing at time k. The track whose root node corresponds to the association with the detection $z_1(k)$ in Figure 7.5 is such an example. A new track is validated in this case. In contrast, the tree whose root node associates with the detection $z_2(k)$ is not selected by $\Theta^{\star k+d}$, so the creation of a new track with the detection $z_2(k)$ is rejected and the whole tree discarded. These two examples show how target appearance events are automatically managed by the tracking procedure. In a similar way, target disappearance events are automatically detected by the selection of a terminated leaf.
After the maintenance operation the best tracking hypothesis $\Theta^{\star k-1}$ is updated with the track from the selected solution $\Theta^\star(k)$ and the root nodes of trees are shifted to the time k since no more ambiguities and incompatibilities remain at this time. The whole tracking procedure is then repeated for the next frame $k+1$: potential tracks are formed, trees are clustered, the best tracking hypothesis is selected, and the trees are updated.

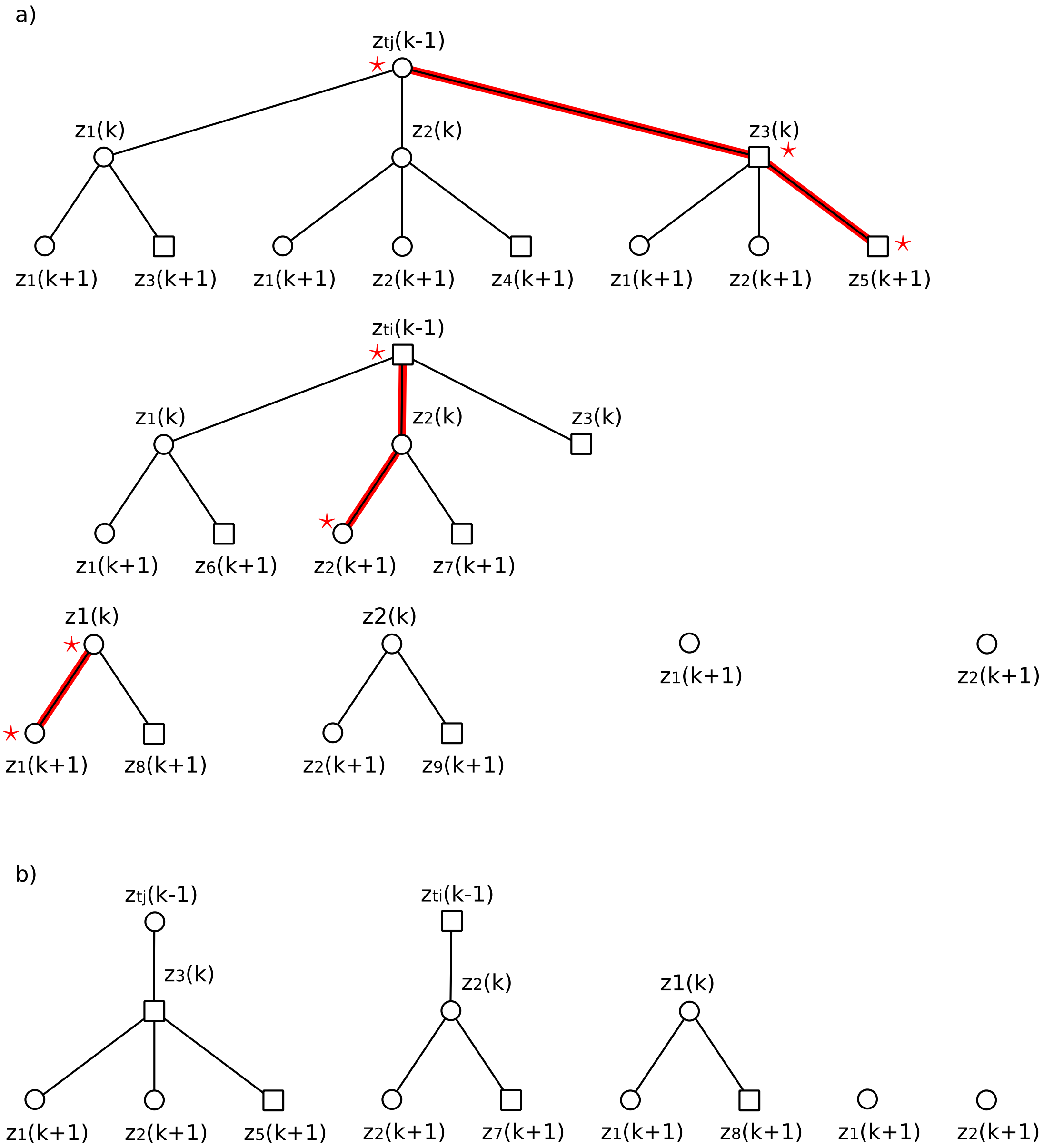


FIG. 7.5: Tracks trees maintenance after association selection. a) The tree representation of the tracking problem presented in Figure 7.3. We indicate with a star the nodes forming $\Theta^{\star k+1}$ and we show in red the corresponding branches in the trees. b) The set of trees is pruned to account for the selection of the solution $\Theta^{\star k+1}$. At time k only nodes selected by $\Theta^{\star k+1}$ remain, hence a number of subtrees are completely discarded, resulting in considerable reduction of the number of leaves. In this example a new track is created whose root node corresponds to the detection $z_1(k)$, while creating a new track with the detection $z_2(k)$ is rejected. The next step is to shift the root nodes to the time k since no ambiguity of association between tracks and detections remain at this time.

7.6 E-MHT characterization and validation

We present next a series of experiments which permit the deep understanding of how the proposed MHT algorithm works. Characterizing the E-MHT and assessing the tracking performance is however complicated by the natural variability of the biological data and the difficulty to measure the quality of the resulting trajectories. We thus introduce first a dedicated framework for quantifying the performance tracking techniques for various imaging conditions.

7.6.1 Methods of evaluation

Synthetic data generation

The benefits of processing synthetic data for particle tracking validation are mainly twofolds : 1) ground truth data is available in a short time as opposite to manual annotation of real images which is extremely time consuming, 2) a wealth of biological and imaging conditions can be simulated without the need of designing and executing different experimental protocols. The quality of synthetic data sets for the purpose of evaluating tracking algorithm however strongly relies on their fidelity to real biological systems and acquisitions devices. For generating synthetic images we distinguish two steps : biological system simulation and images generation, which are assumed independent. For both stages a special care has been paid to the fidelity to real systems. We first simulate systems of particles diffusing in 3D and 2D viscous media. The motion of particles in these types of systems has been described by a number of principles from physics of particles (Chapter 5), and we simulate particle trajectories with respect to these principles. More specifically, the movements of freely diffusing particles between discrete time points are simulated by sampling the subsequent positions of each particle in a multidimensional normal distribution (Eq. 5.18) centered on the previous position with a covariance matrix : $2D\delta t\mathbf{I}_n$ (with n the number of spatial dimensions of the system). The mean of the normal distributions can alternatively be shifted if a force field is simulated (Eq. 5.21).
We also model the possible interactions of particles with cytoskeletal elements by defining for each particle a fixed probability P_{on} to bind to a cytoskeletal structure during the time interval δt between two subsequent frames. When a binding event occurs we simulate a strong and local force field $\mathbf{G}$ applied to the particle. The direction of the force field and its norm are both uniformly sampled in fixed intervals of feasible values. Bound particles thus exhibit a mixture of directed and diffusive motion, while they unbind with a probability P_{off} at each discrete time point. Thus, in accordance with the conveyor belt motion model (Fig. 5.2), the simulated particle can undergo a succession of Brownian motion and mixed diffusion/directed motion, the length of stage is random and linked to the couple of parameters $\{P_{on}, P_{off}\}$.

For each simulation we do not impose a fixed number of particles in the images. Instead, we sample in each frame the number of new targets in a Poisson distribution of fixed mean λ_{NT}. The positions of new particles are then uniformly sampled in the whole current stack. Moreover, each simulated particle can disappear between two subsequent frames with a fixed probability. By doing so the simulated sequences will show a random number of moving particles, some of which appear and disappear over time, whereby the capability of tracking algorithms to account for target appearance/disappearance events can be assessed. We show in Figure 7.6 one example of such 3D simulated trajectories.

After creating the synthetic trajectories at discrete time points we generate the corresponding stacks according to the MPG image model (Eq. 1.1). We first create for each voxel the mean intensity $\lambda(x, y, z)$ of the Poisson term : $\lambda(x, y, z)$ is calculated as the sum of the intensity generated by the particles at this location and of a background intensity which models the

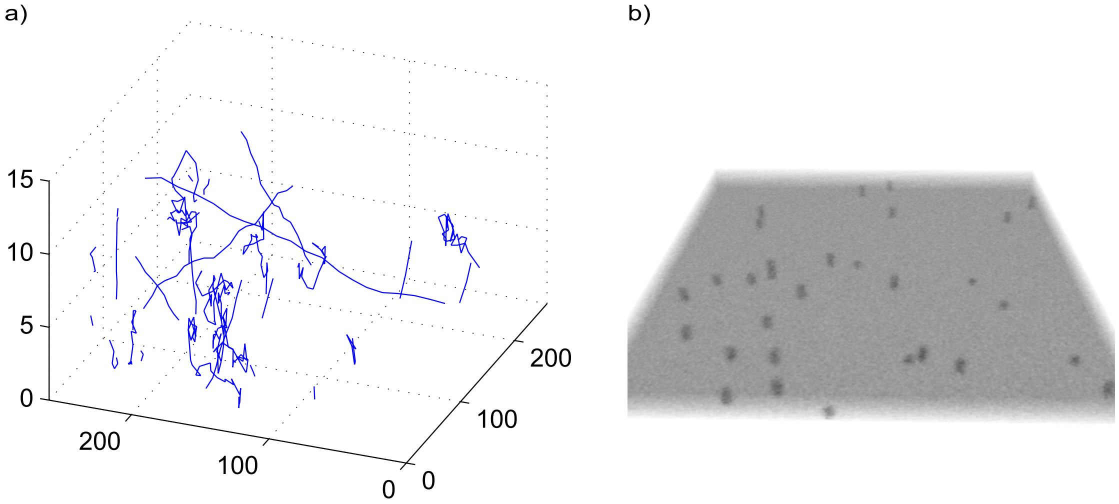


Fig. 7.6: Synthetic data generation. a) Curves linking the subsequent target positions for a number of simulated 3D trajectories. Some trajectories correspond to freely diffusing targets, while others simulate the conveyor-belt motion by alternating between free diffusion and directed and fast motion. Trajectories have different length and different time of appearance, hence modeling target appearance/disappearance events. b) 3D view of a synthetic stack corresponding to one time point of the sequence of images generated for the trajectories in a)

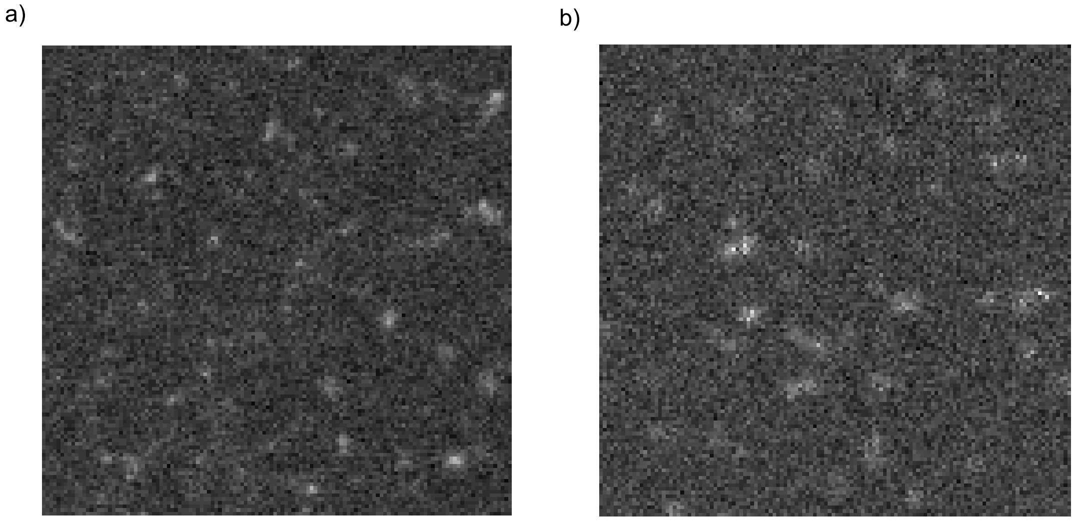


Fig. 7.7: Simulation of images of Golgi units in confocal microscopy. a) Real confocal microscopy of fluorescently labeled Golgi units in Hela cells. b) Simulated image obtained with the proposed procedure for synthetic data generation. The close resemblance between the two images validates the use of the synthetic data set for performance assessment.

contribution of out-of-focus objects, the environment auto-fluorescence, and the dark noise (see Chapter 1). As the intensity profile of the particles we use a theoretical PSF for the CLSM with a fixed amplitude, which is more realistic than the Gaussian model. For each voxel we then sample an intensity value in the Poisson distribution of mean $\lambda(x, y, z)$ and we multiply this value by the gain α. We then simulate the effect of the electronic noise by adding to each voxel intensity one value sampled from the normal distribution $\mathcal{N}(0, \sigma^2)$. In Figure 7.7 we show one real image of fluorescent Golgi units in Hela cells and one simulated image created by using the proposed procedure. The two images are very alike, and discriminating them is difficult without a good knowledge of the real data set. The images we generate have a high fidelity to real biological images, which validates their use for assessing the performance of particle tracking algorithms. The image generation procedure offers a simple way to control the SNR by fixing the intensity amplitude of the particles and the noise characteristics. This flexibility will allow us to study the behavior of various tracking approaches in various imaging conditions.

Measuring the tracking quality with the Jaccard similarity index

Quantifying the quality of an ensemble of tracks with respect to ground truth trajectories is an issue that has not yet found a universal solution. Defining such a tool is indeed made difficult by numerous complicating factors, the most prominent ones being :

- the inaccuracy of the localization process which makes matching detections and particle positions ambiguous,
- the possibility for a true trajectory to be represented by a succession of unlinked parts corresponding to several tracks,
- the fact that tracks corresponding to one particle for a number of frames can switch to another particle due to association errors,
- there might exist tracks corresponding to no real particle but false detections.

To solve these issues we proceed in two steps : first we label the tracks, and then we compute a similarity index between labeled tracks and ground truth trajectories.

We first label each track with regard to the number of true particle positions it contains. One track is said *true* if the positions it gives are localized at a distance to the positions of one single ground truth trajectory which is inferior to a maximal distance d_{max}. Otherwise the track is labeled as being *false*. False tracks are typically composed of a high number of false detections and detections originating from different particles. We note FP (for false positive) their number. In practice we allow few detections to exceed the distance d_{max} : we fix a rate of acceptance α with respect to the required track accuracy. In very low SNR conditions the high number of false and missed detections makes recovering all the target positions extremely difficult. We have experimented different values of α in such a conditions and we have found that relaxing the rate α as low as 0.75 yields satisfactory track labels.

In practice one reference trajectory may be only partially matched by one true track, or it may be matched by several small tracks, or not matched at all. Thus, the rate of true and false tracks is not a good enough measure for assessing tracking performance and we propose to label the ground truth trajectories with respect to the true tracks that match them. A ground truth trajectory is said *recovered* if a single true track matches a significant number of its positions (according to the maximal distance d_{max}). Here again we fix a rate β of positions to be matched by one single true track for the trajectory to be labeled as recovered. A typical setting is $\beta = 0.75$. We note thereafter TP (true positive) the number of recovered trajectories and FN (false negative) the number of non-recovered ground truth trajectories. We aggregate the measures FP, TP and FN in a single measure noted JSC :

$$JSC \triangleq \frac{TP}{FP + TP + FN}, \tag{7.26}$$

which is very similar to the Jaccard similarity coefficient for comparing the similarity and diversity of sample sets [92]. The proposed measure of tracking quality takes values in the interval $[0, 1]$ with the following extreme cases :

- $JSC \rightarrow 1$ when $FN \rightarrow 0$ and $FP \rightarrow 0$,
- $JSC \rightarrow 0$ when $TP \rightarrow 0$ or $FP \rightarrow \infty$.

The set of outputted tracks has therefore a JSC of 1 when every ground truth track is recovered and when no false track is built. In contrast, missing reference tracks or building false tracks both decrease the JSC towards 0.

Using a single track quality index, such as the JSC, aggregating several measures is very useful for comparing different tracking methods in a fair way. However, characterizing further the tracking results requires more specific measures. We thus introduce two supplementary criteria : the true positive rate (TPR) and the recovery rate (RR), which are computed as follows :

$$TPR \triangleq \frac{TP}{TP+FP}, \quad RR \triangleq \frac{TP}{TP+FN}. \tag{7.27}$$

The TPR characterizes the robustness of the tracking approach to false associations, while the RR quantifies its ability to recover a high number of ground truth trajectories. In the following we will say that an approach which results in high TPR tracks is *robust*, while a *sensitive* tracking procedure provides results with a high RR value.

It is worth pointing out that the JSC can be obtained by combining the TPR and the RR as :

$$JSC = \frac{1}{TPR^{-1} + RR^{-1} - 1}. \tag{7.28}$$

Hence, tracking approaches which provide high JSC need to be both robust and sensitive. Equation 7.28 also shows that track sets with similar JSC can have very different values of TPR and RR : one high TPR value can compensate for a low RR, which yields a similar score to that of a track set with a high RR but a low TPR. Hence, in practice, a robust and low sensitivity method can have the same score than a sensitive tracking approach which is not robust. The TPR and the RR will allow us to discriminate between these cases.

An alternative to the JSC is the Dice's coefficient which is often used in information retrieval [79], and can be expressed as :

$$DICE \triangleq \frac{2TP}{(TP+FP)+(TP+FN)} = \frac{1}{2(TPR^{-1}+RR^{-1})}, \tag{7.29}$$

in the tracking framework case. The right-hand side of the Equation 7.29 shows that monotonic transformations between the JSC and the DICE can be easily established on the interval $[0,1]$, hence using either the JSC or the DICE yields the same conclusions when comparing and ranking tracking results.

7.6.2 E-MHT implementation characteristics

Before evaluating the performance of the E-MHT algorithm for various data sets, we first investigate the characteristics of the implementation we have proposed on a single synthetic data set shown in Figure 7.8. It consists in a 3D image sequence of length 20, the size of each stack being $256 \times 256 \times 30$ voxels. One hundred particles are uniformly distributed in the first stack and then freely diffuse. We detect them in the whole sequence with a wavelet-based detector, and then process the set of potential positions with the E-MHT to build the particle trajectories.

Processing time

We first process the synthetic data set with the E-MHT for five different settings of the depth d : 1,2,3,4 and 5. We used a Sun Ultra 40 workstation (setup 1 in Annex) for the experiments

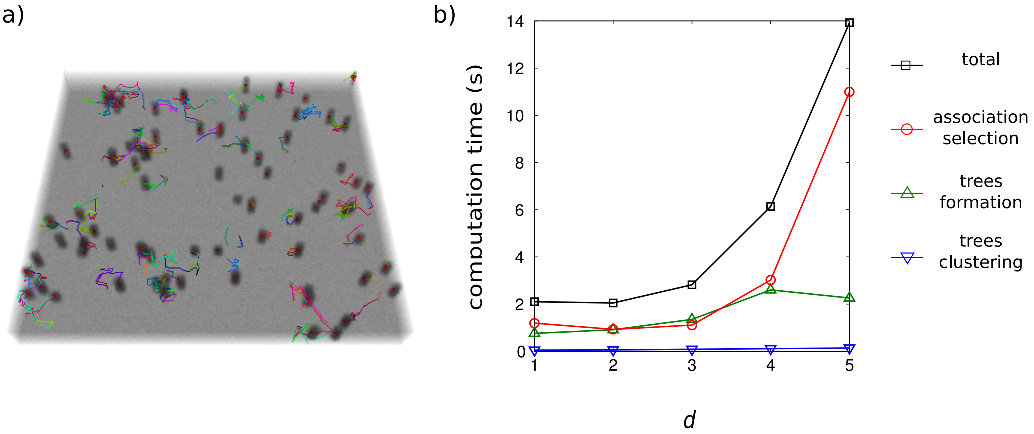


Fig. 7.8: E-MHT computation time for one 3D+T synthetic data set. a) The reference trajectories superimposed on the final time point stack. b) The processing time of the E-MHT as a function of the parameter d. We give the computation time for each stage of the E-MHT. We see that the computation cost increases exponentially with the depth d, but still it remains below 1 second per frame in average. The different curves show that most of the computation time is spent for the selection of the best association hypothesis for $d = 5$. The cost of other stages also increases with d, but in a much slower way.

and we averaged the results of 5 runs for each setting. We plot in Figure 7.8 the estimated computation time as a function of d. We note that the computation time increases exponentially with the value of d, as for the original MHT implementation. However, the processing time still remains very short : in the worst case less than 1 second is spent per frame. We also show in Figure 7.8 the repartition of the computation time between the different stages of the E-MHT. For low d settings the construction of the tracks trees and search for the optimal association are the most time consuming procedures. The required time for the two steps is nearly equal. In contrast, for $d = 5$ a significant increase of the time spent to build the optimal solution is observed, while the time required to build the tracks trees grows slowly with the value of d.
We will show next that the computational cost, which increases exponentially with d, is not an obstacle to the processing of real biological images since using d values larger than 5 is not required for most practical cases.

Parallel computing evaluation

We propose with a second experiment to assess the degree of parallelization of the proposed MHT implementation, and to check its ability to exploit multiple computing units. To do so we have applied the E-MHT to a complex 3D+t synthetic data set containing a high level of noise. We have run the test on a Mac Pro Quad 2.66 GHz (setup 2 in Annex) which offers the possibility to inactivate several cores among the 8 available computing units. We show in Figure 7.9 the measured computation time for 1, 2, 4, 6 and 8 activated cores. The computation time is shown to decrease significantly with the number of cores used, so the parallel implementation is indeed a significant benefit of the E-MHT implementation.
Designing a parallel algorithm for solving the MHT problem is a major challenge because parallel tasks are highly concurrent : they explore and combine the tracks trees simultaneously. A high degree of concurrency between parallel tasks may result in a significant loss of speed because some tasks may be waiting while a needed resource is locked by another running process. In contrast, when the concurrency is minimized the total computational cost is homogeneously

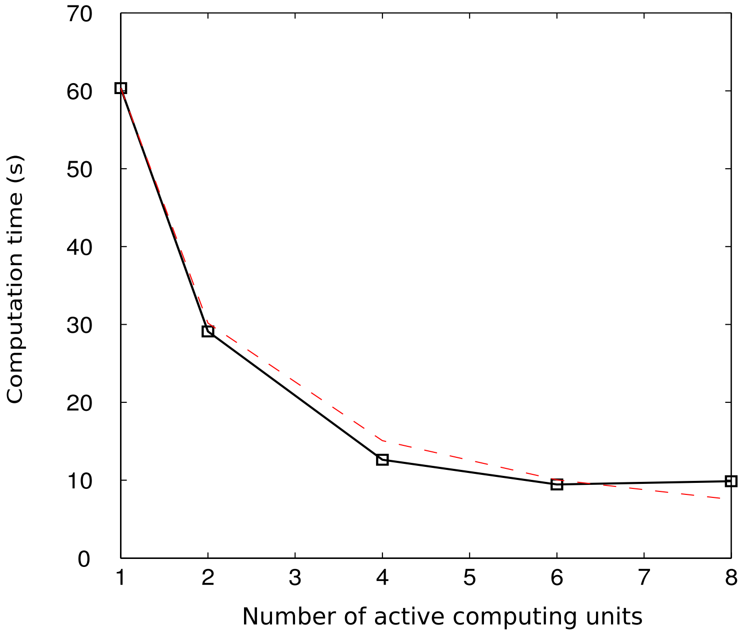


FIG. 7.9: Computation time as a function of the number of computing units. The black boxes indicate the measured time spent by the E-MHT to solve the tracking problem for each configuration. The dashed and red curve shows the predicted computation time when no concurrency exists between parallel processes. The goodness of fit of the theoretical curve shows that the proposed E-MHT implementation minimizes the concurrency between parallel tasks, hence fully exploiting the power of the multi-core computing units.

shared between cores. So required total time can be divided by the number of cores at best. The ideal law for the computational time for a non concurrent parallel algorithm is therefore : $computation\ time = T_1/number\ of\ cores$, where T_1 is the required time when the algorithm is not parallelized (1 enabled core in our case). We show in Figure 7.9 that our implementation achieves results that are close to this ideal law (plotted in red as a dashed line). The goodness of fit of the theoretical law shows that our implementation efficiently minimizes the concurrency between parallel tasks, and thus that the algorithm fully takes advantage of the power of multi-core computation units. Multi-core processors will probably widespread in future years, and the number of cores is planned to increase for every processor, so, we expect the E-MHT to take full advantage of these technological developments.

7.7 Benchmarking of particle tracking methods performance in poor imaging conditions

We have shown in Part I that a number of solutions have been proposed for tracking particles in fluorescence images, but, to the best of our knowledge, no systematic comparison of these approaches has been done and published. Such a study would however be of great interest to identify the pros and cons for each method. We thus present next a series of three experiments performed on both synthetic and real data sets, in 2D and 3D, with a number of state-of-the-art techniques and with the E-MHT.

We have already mentioned in Part I that handling poor imaging conditions is a major challenge when tracking particles in biological environments. We have thus designed a benchmark for assessing the capability of algorithms to track particles in low SNR fluorescence images. These tests allow assessing the characteristics of the different approaches for various conditions, and prove that the novel tracking algorithm we propose solve some limitations which are inherent to other existing solutions.

7.7.1 Low SNR synthetic images processing

Benchmark setup

We have simulated 2D sequence of images since most state-of-the-art methods have been designed for 2D+t data sets. The 3D+t case will be investigated with the third experiment giving results for real 3D+t data processing. For the benchmark we have generated some synthetic data sets using the procedure presented in 7.6.1 with the following settings :

- 256 × 256 pixels images with a background of intensity 50 and corrupted by a Gaussian noise with a variance $\sigma^2 = 25$,
- sequences of length 50 frames,
- 20 particles exhibiting a conveyor-belt motion (random transitions between free diffusion and directed and fast motion) are present in the first frame,
- the mean number of particle appearance events is 0.1 for one frame,
- each particle can disappear between subsequent frames with probability 0.01.

We show in Figure 7.10 two typical examples of trajectories generated with these settings, which illustrate the variability of the particle movements and the possibility for targets to appear and disappear over time. In order to assess tracking performance in images of various quality we have chosen five different settings for the amplitude of the intensity profile of particles : 30, 20, 15, 10, 8. These settings corresponds to a regular decrease of the SNR condition : $SNR = 2.93$, 2.05, 1.58, 1.08 and 0.88, respectively. As shown in Figure 7.10 for the highest SNR conditions particles are significantly brighter than the background, however, for low SNR conditions the particle intensity is so low that they are barely visible. The SNR condition will thus have deep impact on the quality of the sets of detections used by tracking algorithms. The ability of tracking approaches to deal with corrupted detections will thus play a key role for the tracking performance in low SNR conditions.

In this study we have compared the results obtained with five state-of-the-art tracking methods (see Part I) and the E-MHT :

1. **IML**[67] : probabilistic tracking embedding different motion models, and for which the association decision is instantaneous, in contrast with the MHT approach. The particle detector is the wavelet-based method presented in [129]. We have run a Java implementation of the IML algorithm.

2. **GC** [94] : probabilistic tracking with a linear motion model and the ability to cap gaps between partial tracks. We propose to test two different detectors : the detection process proposed in [94] which relies on matching Gaussian shapes in the images, and localizing spots with the wavelet-based detector [129]. We name these two approaches **GC1** and **GC2**, respectively. The GC algorithm is implemented as a Matlab software with some routines compiled from C language sources.

3. **FPT** [147] : the Feature Point Tracking algorithm is not probabilistic approach, but is distance-based instead. The FPT is implemented as a plugin for the ImageJ software and is provided with its own particle detection procedure which consists in smoothing the image with a Gaussian kernel and analyzing the resulting histogram of intensity values [49].

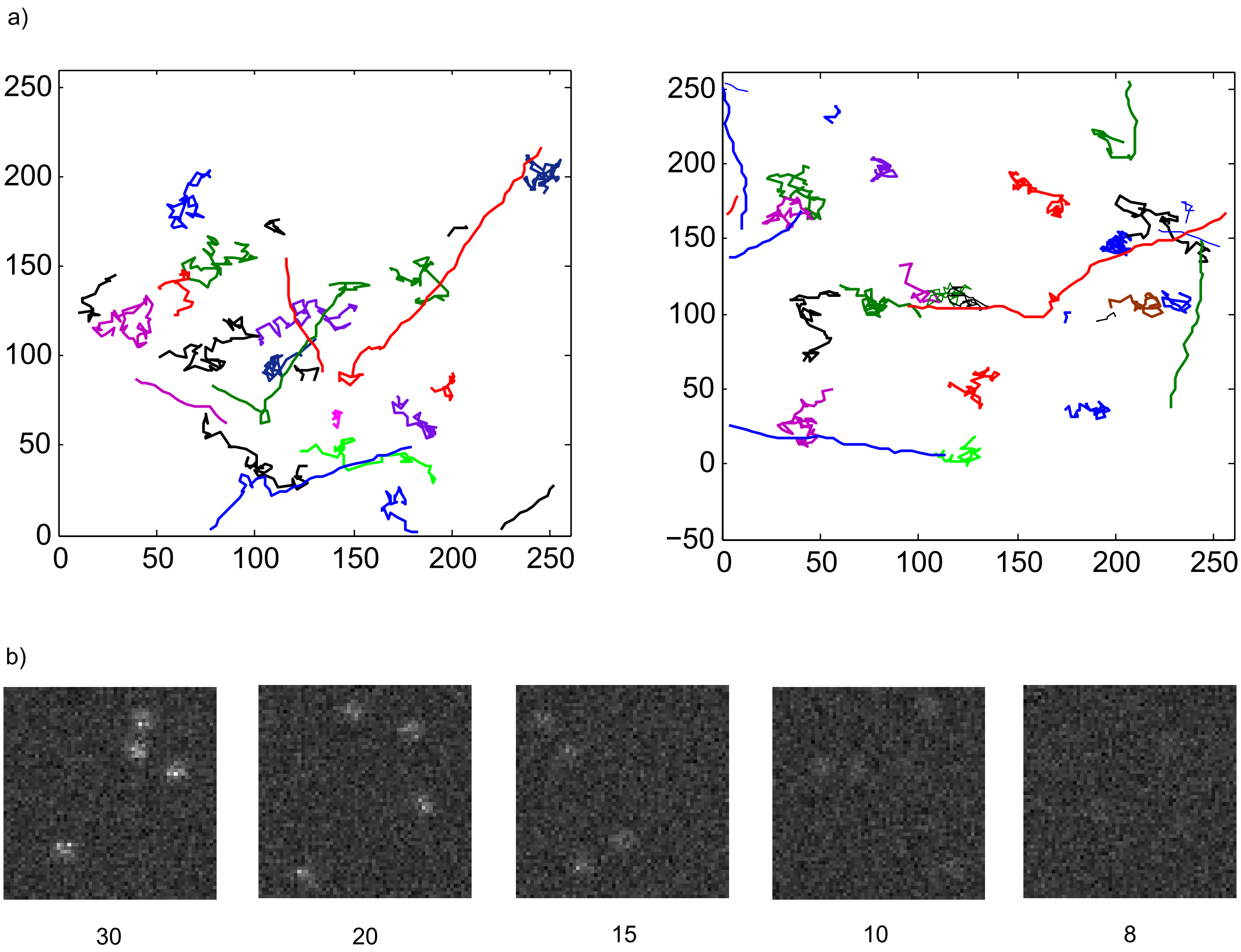


FIG. 7.10: Tracking benchmark for assessing tracking performance in different imaging conditions. a) Two set of trajectories generated for the benchmark. Most simulated particles are freely diffusing, while some of them suddenly exhibit a fast directed motion. b) A series of cropped images around some particles with different intensity amplitudes. From left to right the amplitude values are : 30, 20, 15, 10 and 8. For amplitudes values below 10 discriminating particles from noise is very ambiguous.

4. **E-MHT** : the proposed MHT algorithm with $d = 4$ when incorporating multiple models of motion. We use the detections provided by the wavelet-based particle detector.

For each SNR condition 10 image sequences have been generated and processed with these five techniques. We have adjusted the settings of the detection step and the tracking procedure to maximize the JSC of the resulted tracks for each image quality condition. We provide in Table 7.1 the complete set of results : mean JSC, RR and TPR for each method in every imaging condition. We give the corresponding curves for the JSC, RR and TPR mean values in function of the SNR in Figure 7.11.

Results

We first note that the E-MHT provides results with much higher JSC values than other techniques for every SNR condition. Indeed, the second best observed JSC value is obtained by the FPT in the best imaging condition (JSC=0.50), but still in this case the difference with the E-MHT results (JSC=0.82) is large. In low SNR conditions the difference between the E-MHT

	SNR	**2.93**	**2.05**	**1.58**	**1.08**	**0.88**
FPT	TPR	0.91	0.48	0.27	0.03	0.01
	RR	0.52	0.52	0.38	0.12	0.01
	JSC	**0.50**	**0.30**	**0.17**	**0.02**	**0.00**
IML	TPR	0.43	0.34	0.29	0.23	0.20
	RR	0.86	0.82	0.70	0.61	0.50
	JSC	**0.40**	**0.32**	**0.26**	**0.20**	**0.17**
GC1	TPR	0.86	0.90	0.84	0.65	0.09
	RR	0.44	0.46	0.33	0.18	0.001
	JSC	**0.42**	**0.44**	**0.31**	**0.17**	**0.01**
GC2	TPR	0.71	0.61	0.54	0.52	0.29
	RR	0.44	0.41	0.35	0.31	0.18
	JSC	**0.38**	**0.32**	**0.27**	**0.24**	**0.13**
E-MHT	TPR	0.95	0.93	0.90	0.90	0.88
	RR	0.86	0.92	0.86	0.81	0.69
	JSC	**0.82**	**0.86**	**0.79**	**0.75**	**0.63**

TAB. 7.1: Benchmark results for tracking in low SNR conditions : 4 state-of-the-art tracking techniques and the proposed E-MHT algorithm are tested. The results are averaged over 10 synthetic sequences generated with similar settings.

and other techniques grows significantly since the JSC values for state-of-the-art methods quickly decrease towards zero. In contrast, the quality of the tracks built by the E-MHT is remarkably stable with a low sensitivity to the image quality : the maximum JSC value is 0.86 for SNR=2.05, while JSC is still 0.63 in the worst conditions (SNR=0.88).

The FPT algorithm provides the best results among state-of-the-art tracking techniques in high SNR conditions, however, the JSC scores dramatically drops down with the decrease of the image quality. The TPR and RR curves show that the most prominent cause of the JSC decrease is a sudden increase of the number of false tracks when the SNR is lowered (the TPR decreases much more rapidly than the RR). These numerous false tracks originate from the association of tracks with false detections which are dense in low SNR conditions and indicate that the detector used in FPT is not robust. Associating true tracks to false detections also degrades the RR. Thus, the FPT is not robust enough and should be used only when the image quality is good (SNR>3).

The IML algorithm provides results with a remarkably good RR even with low SNR conditions. This ability to recover numerous reference tracks is partly due to the capability of the IML to deal with various motion types thanks to the embedded IMM filter. Building complete tracks is also facilitated by the small number of missed detections yielded by sensitive settings of the particle detector. These permissive settings however result in a high number of false detections which are not handled in a proper way by the IML algorithm, as shown by the low TPR results. In practice the algorithm does not discriminate between proper and spurious detections, so the number of false tracks is very high and the JSC scores are poor.

The GC1 and GC2 algorithms both give JSC results similar to the IML : a slow decrease of JSC is observed from approximately 0.40 to values below 0.20 for the worst image conditions. The TPR and RR values however indicate that the three procedures exhibit very different behaviors. Contrary to the IML, the GC1 approach gives robust results in high SNR conditions, but recovers only few reference trajectories. We explain these differences mainly by the use of a very

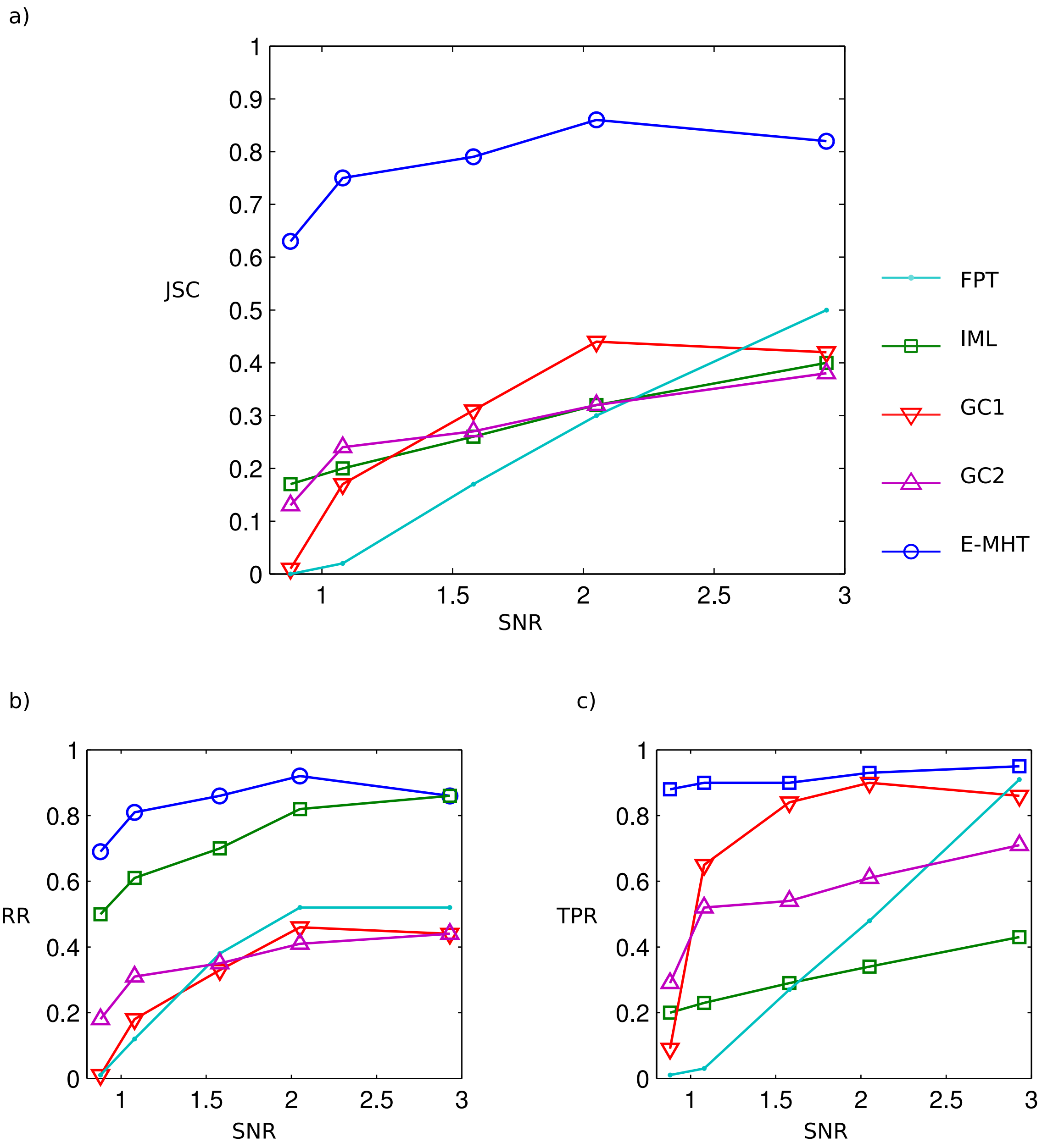


FIG. 7.11: Tracking in low SNR conditions benchmark results.

different particle detection technique : in the GC1 the detection procedure is based on matching Gaussian functions in the images and detecting significant matches. This approach is robust to artifacts, so, it produces only few false detections, hence yielding tracks with high TPR values. The detection technique is however unable to detect low intensity particles, so only few reference trajectories can be recovered in low SNR conditions. When changing the detection method to the wavelet-based technique the robustness of the results is significantly degraded by the presence of numerous spurious detections, as shown by the decrease of the TPR values between GC1 and GC2. Interestingly, the RR curves for GC1 and GC2 show that the number of true tracks is not significantly increased in this case. The high number of reference tracks which are missed may be explained by the inability of the GC to model accurately the different particle motion types with only one motion filter.

The study of the TPR and RR curves shows that the high JSC values for the E-MHT are obtained thanks to both very high TPR and RR values. The slight decrease of the JSC values for low SNR conditions is mainly due to a decrease of the number of true tracks built : the RR drops from 0.92 (SNR=2.05) to 0.69 (SNR=0.88), while the largest TPR loss is only 0.07. The E-MHT is therefore a very robust tracking technique : it builds only very few wrong tracks whatever the imaging conditions. In low SNR conditions fewer true tracks are built, but still with a very high accuracy. This is important for downstream analysis because a high confidence can thus be given to the outputted tracks, whatever the conditions.

The slight decrease of the RR in low SNR conditions is mainly due to the definition of the the score function which is optimized by the E-MHT algorithm : maximizing $\mathcal{L}(\Theta^l)$ consists in making the false detections and the particle models compete. However, in low SNR conditions the likelihood of the false detections model increases, so events with low probability are more likely to originate from noise than particles, and fewer tracks are built by the E-MHT. The extent of the RR value decrease is nevertheless small compared to the one of other tracking techniques. The reasons of this more satisfactory behavior are twofolds : an accurate modeling of the target existence, and a better integration of the temporal information by the E-MHT. Indeed, at each frame we account for all the past and $d+1$ future frames, hence discriminating particles from noise is still feasible, despite the increased likelihood of the false detections model in low SNR conditions.

The RR results also show that the E-MHT and the IML are able to recover significantly more reference tracks than other techniques : the RR is around 0.8 for the IML and E-MHT in high SNR conditions, while it hardly exceeds 0.4 for other techniques. The shared characteristics of the IML and E-MHT are the use of a sensitive particle detector and the embedding of multiple kinetic filters for modeling the motion of particles. These two characteristics are thus highly desirable when tracking particles in low SNR conditions.

7.7.2 Golgi units tracking in cluttered 2D images

In order to assess and compare the efficiency of the different methods we have used images sequences from a challenging biological assay. We show in Figure 7.12 one 504 by 405 pixels confocal slice taken from a 462 frames DSCM sequence of Golgi units in Chinese Hamster Ovary cells. Golgi units containing the protein appear as bright spots in an uneven background. The tracking task is made challenging by the numerous targets with various motions and the poor quality of the images. Indeed, the PSNR is below 3 for low intensity particles, which makes handling false and missed detections a major issue.

In order to assess tracking performances, an expert has manually identified Golgi unit trajectories in a 72 by 64 pixels crop during 128 time steps. In this small area 214 target trajectories has been labeled. We show in Figure 7.13 few images from the sequence and the corresponding reference tracks

We have applied the E-MHT, IML, GC, FPT. In addition, we have adapted to the Brownian case the MHT algorithm proposed by Cox and Hingorani in the computer vision field [47] and combined it with a wavelet-based detector used for the IML and E-MHT. The depth of this MHT was fixed to 3, and for the E-MHT four d settings were investigated : $d = 2, 3, 4$ and 5. We summarize tracking results obtained by the investigated methods in Table 7.2, while tracks obtained by the E-MHT for $d = 4$ ar shown in Figure 7.13.

Results show that the corruption of the detection set by spurious measurements leads instantaneous association algorithms such as IML and GC to produce many false tracks. It reveals

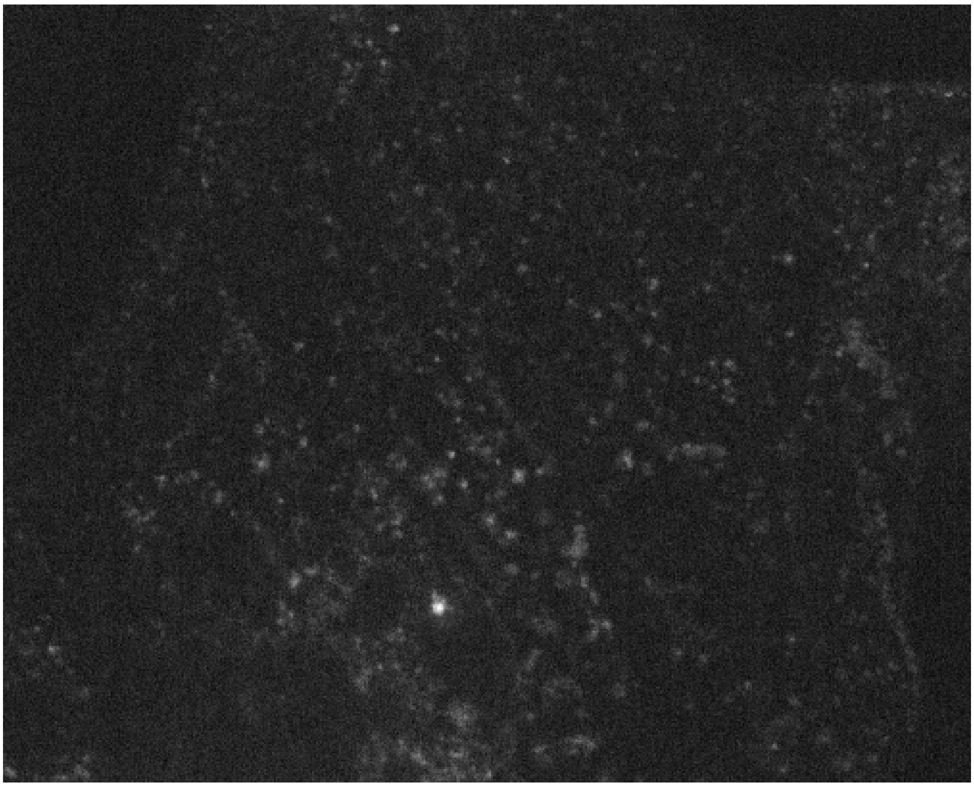


FIG. 7.12: Fluorescently labeled Golgi units in confocal microscopy.

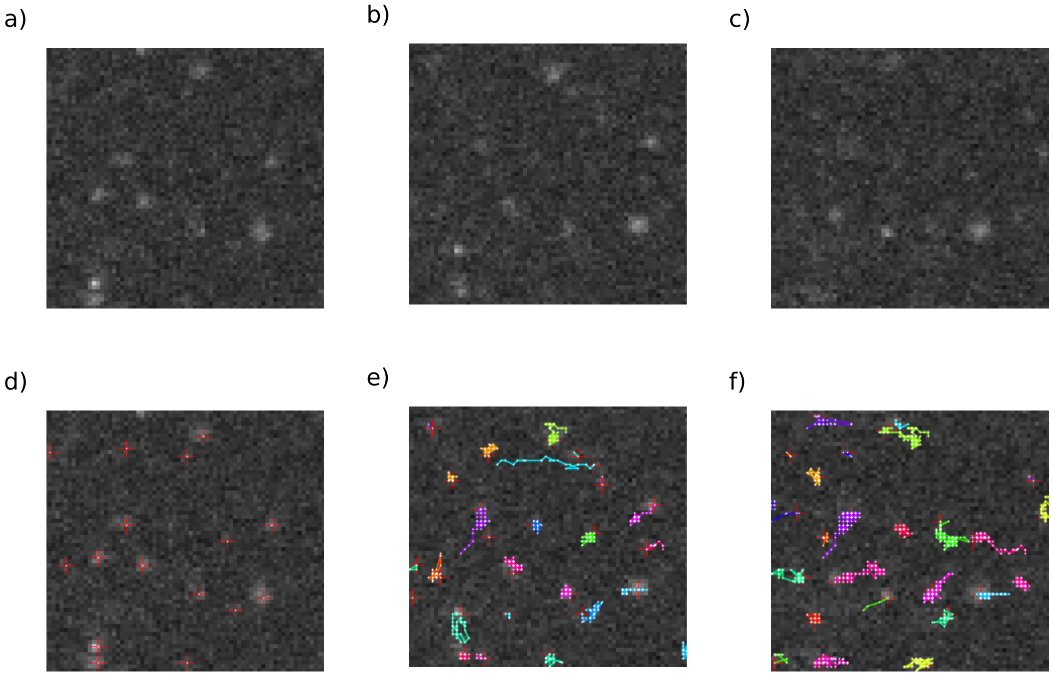


FIG. 7.13: Golgi units in a 2D restricted area at three times (first row). Second row shows the corresponding tracking by the E-MHT.

	TP	FP	TPR	RR	time	**JSC**
GC	51	30	0.63	0.24	6.9s	**0.21**
IML	175	142	0.55	0.82	2.3s	**0.49**
FPT	16	18	0.47	0.07	$<1s$	**0.07**
MHT [47] d =3	191	95	0.67	0.89	25s	**0.62**
E-MHT d=2	183	37	0.83	0.86	8.8s	**0.73**
E-MHT d=3	184	35	0.84	0.86	11.9s	**0.74**
E-MHT d=4	177	34	0.84	0.83	17.0s	**0.71**
E-MHT d=5	173	25	0.87	0.81	25.6s	**0.72**

Tab. 7.2: Golgi units tracking results. Top : performances for standard algorithms. Bottom : performances of the E-MHT technique.

their lack of robustness to cluttered conditions. The high corruption rate of the track set makes the extracted trajectories useless for most downstream analyses. On the other hand, multiple hypothesis tracking algorithms show an improved robustness to spurious detections which is due to the automatic spurious detections exclusion, made possible by the processing of both past and future frames.

When comparing the different MHT techniques, our proposed approach achieves the best JSC results, whatever the value of the depth. The main reasons of the superiority of the E-MHT over the reference MHT implementations are threefold :

1. we incorporate dedicated models of motion for biological particles,
2. the robustness to false detections is improved further by the incorporation of perceivability models, which yields a lower number of false tracks,
3. hypotheses pruning is exact instead of being heuristic as for the MHT of Cox.

It is worth pointing out that the algorithm depth influences the resulting tracks only to a limited extent. For instance, increasing the depth improves only slightly the robustness to false detections. This characteristic is explained by the short range of auto-correlation in diffusive motions, which are specific to biological applications. Indeed, when observing diffusive motions, only the few previous and next frames contain a significant information for the current tracking state. The E-MHT can therefore be used in these applications with a small depth ($d \leq 5$), which reduces the computation time without any loss of performance.

The E-MHT procedure with a depth of 4 was also applied to the whole sequence of images. It took only 6 minutes, with a Mac Pro Quad 2.66 GHz (setup 2 in Annex), to process the 462 frames which is fast when taking into account the complexity of the scene. The 4024 resulting trajectories are displayed in Figure 7.14. It shows the ability of the algorithm to deal with high densities of targets, and to take into account various types of movements; it is worth noting that results obtained by the other methods are not exploitable because of their poor quality. Result movies can be found on line at *http ://bioimageanalysis.org/2435/*.

7.7.3 3D tracking of fluorescent beads

In a third experiment we study the tracking performance of three algorithms : the GC, IML and E-MHT, with the processing of a real 3D+t data set. We study the motion of fluorescent synthetic beads (Fluorescent Fluospheres® from Molecular Probes®) in water which provide a system which is easier to control and annotate manually than when imaging biological objects *in*

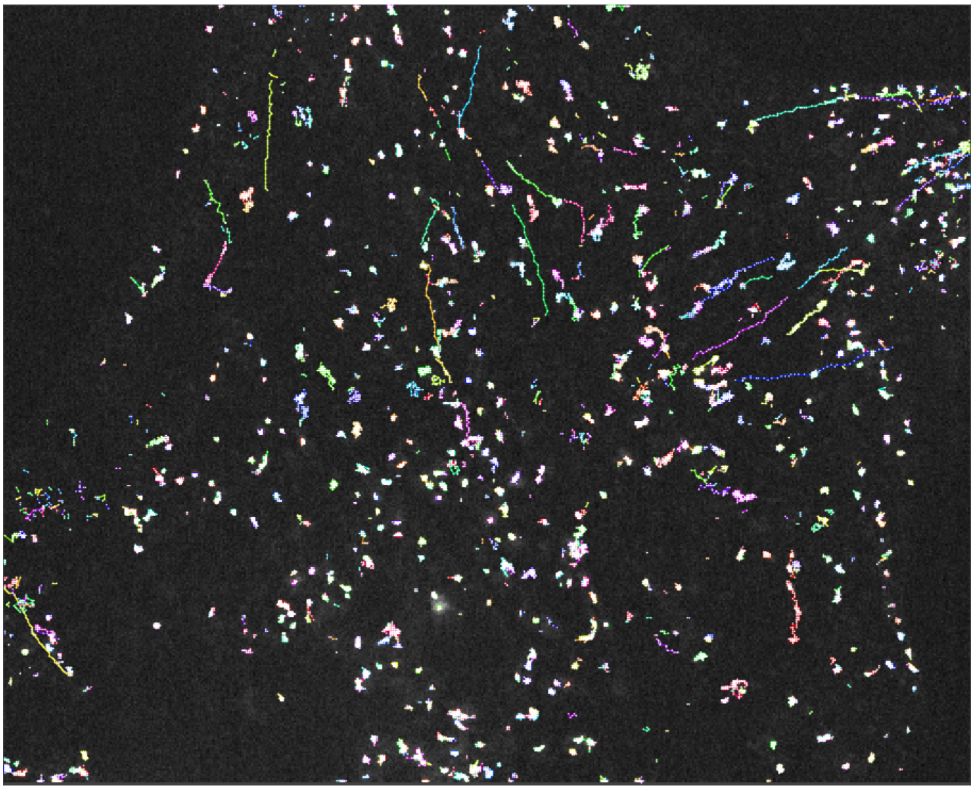


FIG. 7.14: Trajectories of Golgi units built by the E-MHT.

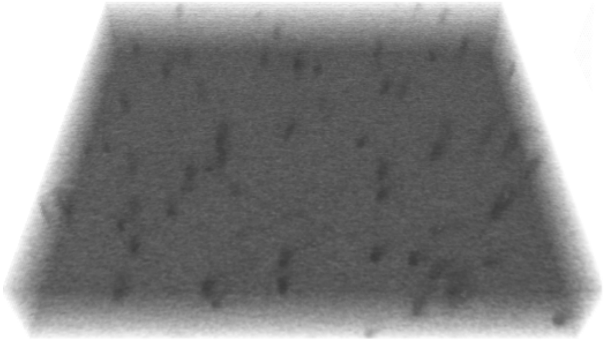


FIG. 7.15: 3D imaging of fluorescent beads diffusing in water.

vivo. We acquired with a DSCM a number of sequences with a duration of 15 s, each one composed of stacks of images containing 21 confocal slices with a spacing of 250 nm. The beads have a diameter of 0.2μm, which is close to the size of many biological objects, and small compared to the resolution of the DSCM, so they appear as blurry and small blobs in the 3D images. As an example, we show in Figure 7.15 one acquired stack when exposing the sample 20 ms to the laser light.

In theses conditions the average SNR is moderate : $PSNR = 33.2$. Some particles have however a low intensity level, so a detector with sensitive settings is required. For 50 manually annotated

	TP	FP	TPR	RR	time	**JSC**
GC	59	18	0.77	0.42	$5s$	**0.38**
IML	108	63	0.63	0.78	$3s$	**0.54**
E-MHT	110	7	0.94	0.80	$7s$	**0.76**

Tab. 7.3: Performance of three methods for tracking fluorescent beads in a 3D sequence of images. The tracks obtained with the three methods are compared to 138 manually labeled trajectories. Top : the tracking results obtained with two state-of-the-art algorithms : the IML and the GC. Bottom line : performance of the proposed E-MHT for $d = 3$.

frames the wavelet-based detector [129] applied in 3D provides 3054 correctly estimated particle positions, while 187 locations are missed and 904 detections are wrong. Some missed positions are due to particles that temporarily aggregate and cannot be resolved by the detector. Dealing with missed and false detections is therefore an important issue when tracking the fluorescent beads.

In water the fluorescent beads diffuse in a very fast manner and the distance between the subsequent positions of particles can be large ($\sim$5 voxels), which makes solving the association problem between tracks and detection ambiguous. As shown in Figure 7.16 the beads move in a non-Brownian way : the trajectories are highly directed due to a flow of water in the sample which is caused by some thermal inhomogeneities and the low viscosity of the medium. Modeling this particular motion in an appropriate way is therefore required to solve the association problem.

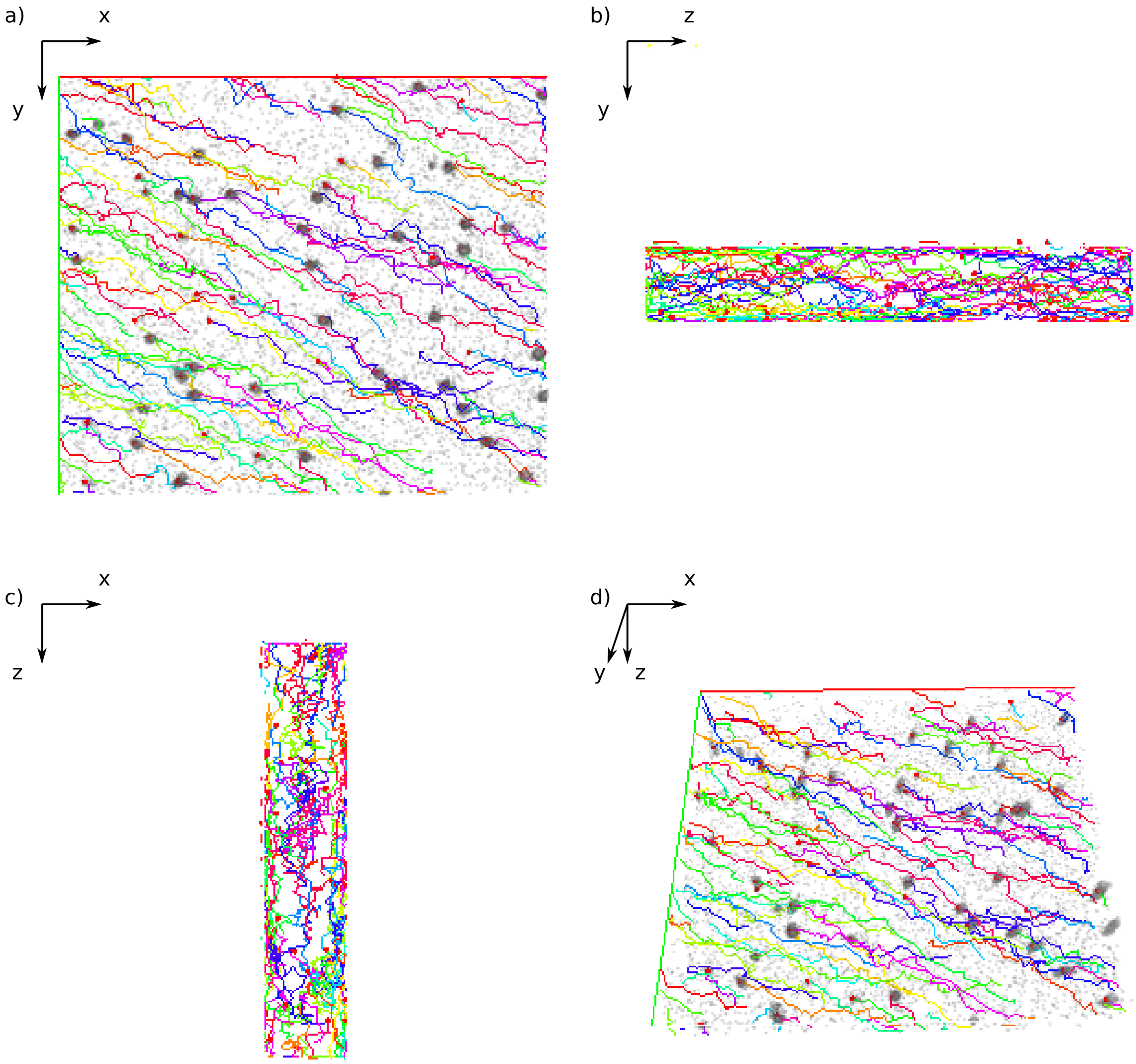


FIG. 7.16: 3D trajectories built by the E-MHT when tracking fluorescent beads in water. Trajectories are superimposed to the final timepoint stack. a) Top view of the trajectories. b) and c) side views. d) 3D view. The tracking results show a very directed movement of particles which can be due to a flow of water.

We have used three tracking algorithms : GC, IML, and the E-MHT, both combined with the wavelet-based spot detector applied in 3D. The GC uses a linear motion model, as well as the IML and E-MHT which embed the motion model for a particle diffusing in a force field (Chap. 5). For the E-MHT we have set $d = 3$. We have annotated 50 frames from a 3D+T image sequence to obtain 138 ground truth trajectories. The results for these three tracking procedures when compared to the reference trajectories are provided in Table 7.3. In summary, we note mainly three points :

- The GC algorithm is quite robust to false detections ($FP = 18$), however, it is not sensitive enough and does not provide a proper motion model. Hence, only few reference trajectories are built ($TP = 59$ for 138 reference tracks).
- The IML is able to recover a high number of ground truth trajectories thanks to a good motion modeling ($TP = 108$) but it is not robust to false detections. Thus, the TPR of the tracks is poor and the JSC is significantly degraded ($JSC = 0.54$).
- The E-MHT is both robust and sensitive, and it provides high quality tracks ($JSC = 0.76$).

These results confirm the first ones we obtained in the 2D case which displayed similar kind of behavior for the three tracking algorithms. The efficiency of the three tracking approaches thus does not significantly depends on the dimension of the data set. Here again we conclude that for tracking real images an accurate modeling of the target behavior is required, as well as a dedicated model of false and missed detections. The E-MHT is the only 3D+T tracking approach we have tested that fulfills all these requirements.

7.8 Conclusion

In this chapter we have described a novel formulation of the MHT algorithm which is dedicated to biological particle tracking in fluorescence images. The E-MHT is based on the probabilistic framework presented in Chapters 5 and 6, thereby it is able to account for different types of particle motion and it takes advantage of the model of target perceivability proposed for particles. Accounting for knowledge from future frames makes the statistical scores for tracking more discriminant. Hence, the E-MHT can decipher challenging tracking problems, such as handling very poor imaging conditions, which cannot be reliably solved when only the "past" is modeled.

In order to achieve the best quality results, the proposed implementation of the MHT does not involve any simplification such as hypothesis merging or heuristic pruning of solutions. This choice to promote the quality of the tracking induces an increased complexity of the tracking problem which has to be dealt with. To do so, we have proposed a MHT scheme that exploits the tree structure of the potential tracks and makes enumerating the possible associations and selecting the best one an extremely efficient process. We use a massive, but exact, association pruning that is based on a *branch and bound* scheme. This technique relies on an association score choice which decreases with the additions of tracks to the association, giving a hard decision threshold for pruning. Moreover our algorithmic design allows massively parallel processing which makes the different steps of the algorithm very fast on parallel computing architectures which have become widespread in recent years.

In order to validate our approach we have proposed a series of experiments with various data sets : synthetic 2D image sequences with different SNRs, fluorescent Golgi units in a dense and cluttered 2D environment, and 3D images of fluorescent beads diffusing in water. For each experiment we have compared the E-MHT to a number of state-of-the-art techniques. The benchmark with synthetic images has revealed that previously proposed solutions are not able to handle poor imaging conditions in a proper way : a number of techniques are too conservative and are thus not sensitive enough, while other ones are not robust and produce many false tracks. In contrast, the E-MHT is both robust and sensitive despite the poor quality of images yielding a high corruption of the detection set. The E-MHT is indeed able to check over time the consistency of the detections by accounting for both past and future frames thanks to the statistical framework. By doing so, false detections are automatically discarded and missing ones are compensated for. The performance results of the tracking algorithms with synthetic data have been confirmed by experiments with real data sets. Hence, the E-MHT has been proved to be able to process in a short time real images containing hundreds of particles, despite the cluttered conditions. The capabilities of this new algorithm therefore open the way to novel biological studies which where until now impossible to analyze with a sufficiently high quality.

Part III

Biological applications

Chapter 8

Application of the proposed methods in biological studies

8.1 Golgi units tracking in Drosophila melanogaster oocytes

We have participated to a biological study focusing on the transport of Golgi units in the ovocyte of the Drosophila melanogaster and their import into the oocyte [126]. Golgi vesicles have been labeled with GFP and a DSCM has been used to acquire numerous image sequences of the ovocytes. In Figure 8.1 we show images acquired for this study, which illustrate the compartimentation of the Drosophila ovocyte and the packing with follicle cells. In this study we were particularly interested in the transfer of proteins between the nurse cells and the oocyte. Nurse cells are indeed known to export proteins to the oocyte during the oocyte growth. These transfers are achieved through canals called *ring canals* and aimed at characterizing the mechanisms supporting this process.
Characterizing the import of Golgi units into the oocyte has required studying their dynamics by establishing their individual trajectories. To do so, we have first applied a wavelet-based detection technique[129] which has flattened the background at the vicinity of the canal which produces a high level of intensity (as shown in Figure 8.1). Because of the low PSNR (~ 5) the detections sets were corrupted by some spurious detections : we measured up to 25% of false detections, and some targets were transiently missed by the localization procedure. We have thus applied the E-MHT technique which was capable of handling such poor imaging conditions. The IMM filtering technique was included in the E-MHT to model the wealth of different motion types that Golgi units may exhibit.
We show in Figure 8.2 an image of Drosophila ovocyte and the 111 trajectories built by the E-MHT technique with a depth of 4 for the corresponding image sequence. Result movies are available on line at http ://bioimageanalysis.org/2436/. The built trajectories revealed various types of motion. In the nurse cell's cytoplasm, two categories of movements were observed : random Brownian movements and fast linear movements interrupted from time to time by short stops or backward shifts. The latter would be compatible with directed transport supported by microtubules. We have also shown that in front of the ring canal Golgi units paused and clustered, exhibiting a Brownian like motion. On the other hand, it was observed that the transit through the ring canal was performed in a directed and fast manner for some selected units, which is incompatible with simple diffusion.
In this study the velocity measurements of targets were investigated : the speed was computed at each time point for directed movements, revealing a switch from fast motion during the

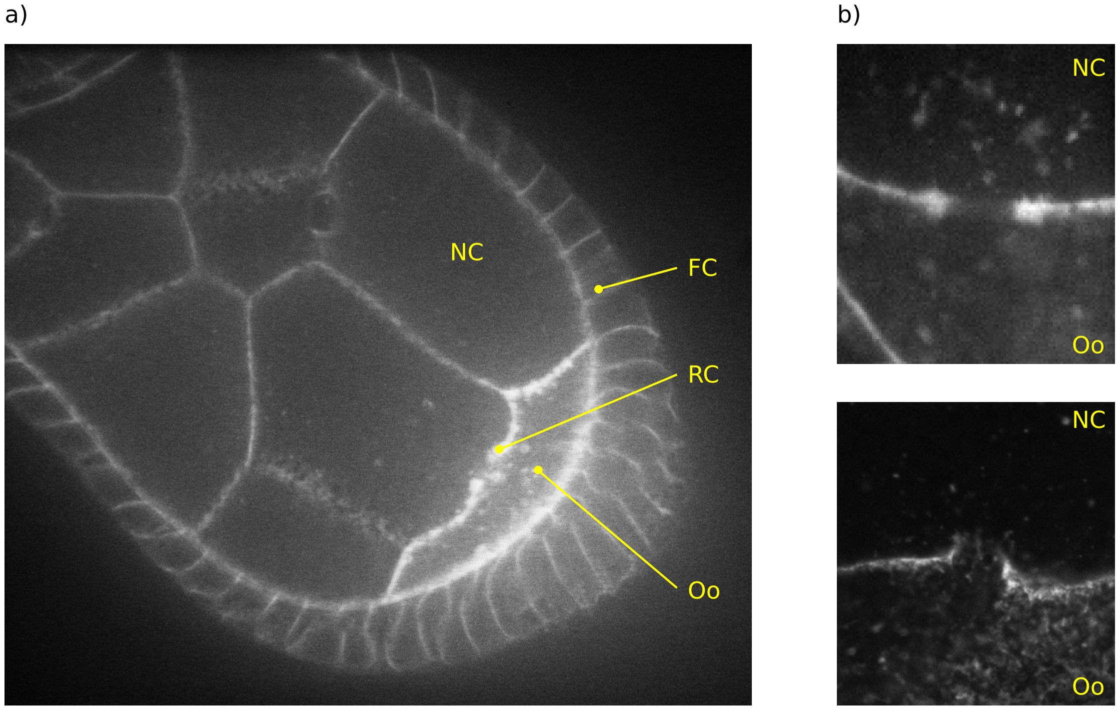


Fig. 8.1: Drosophila melanogaster ovocyte in confocal fluorescence microscopy. a) An ovocyte showing some nurse cells (NC) linked to the oocyte (Oo) by ring canals (RC). The membranes of follicle cells (FC) packing the ovocyte are also visible. b) Two ring canals between nurse cells and oocytes. Golgi units are shown to aggregate near the entry of the ring canal on the nurse cell side, while the density of particles is very high on the oocyte side.

approach (0.190μ m/sec±0.024) of the ring canal to slower movement during the canal transit (0.110μm/sec±0.026) and thus suggesting that different molecular motors might be involved in the ring canal approach and its transit. The mean-squared analysis (MSD) [146] of the trajectories has also permitted discriminating two groups of particles crossing the canal : one with parabolic MSD curves and the second with more linear curves. The first group of units was thus characterized by a high velocity and directed movements, while particles in the second group moved in a more Brownian manner. Interestingly, we show in Figure 8.3 that the motion type matches with the spatial distribution of the trajectories : fast particles have crossed the canal at the center, while slow particles have crossed near the canal rim. Transit in the middle of the ring canal exhibited straight trajectories and a continuous motion, which are features of active transport. On the contrary, it was shown in [126] that Golgi units crossing close to the canal rim displayed more complex trajectories of acceleration and short stops which suggested sequences of Brownian movements and active motion.

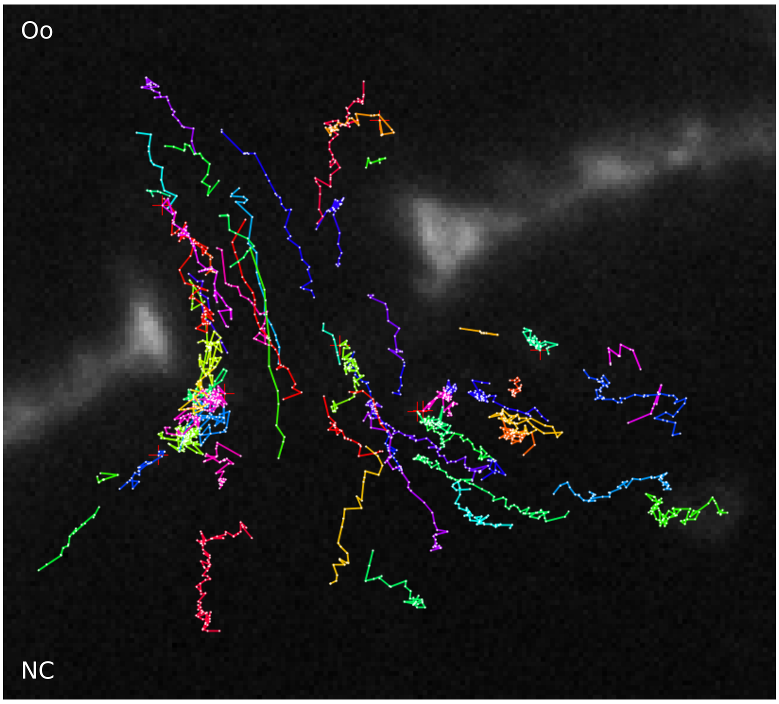


Fig. 8.2: Trajectories of Golgi units built by the E-MHT in DSCM images (NC : nurse cell, Oo : oocyte). A wealth of motion types are observed : in the cytosol of the NC particles head towards the RC, near the canal Golgi units are aggregated, while some of them suddenly cross the canal in a directed and fast way.

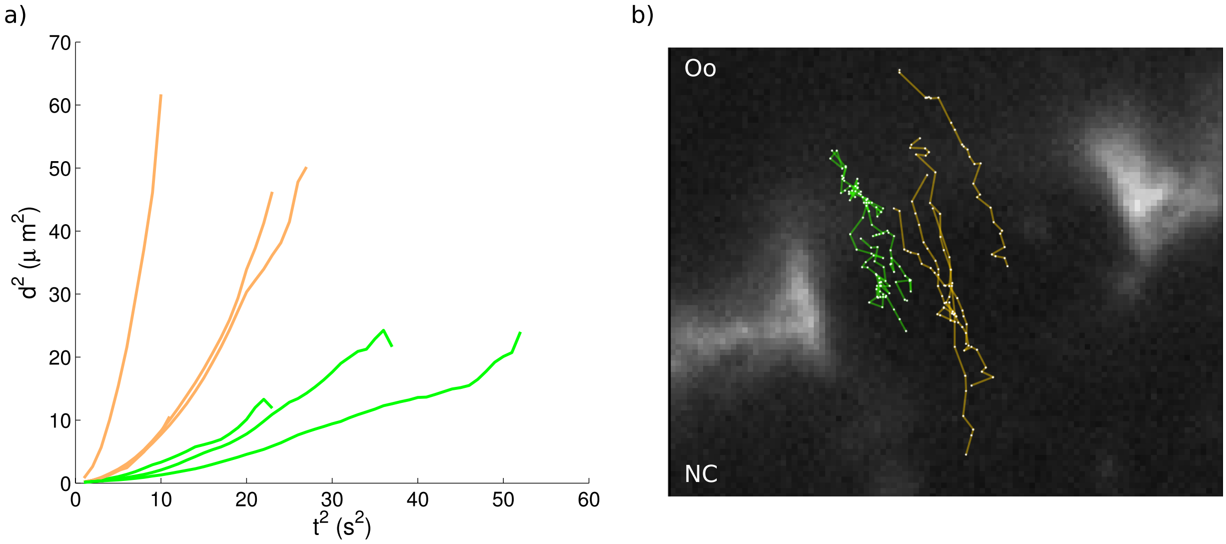


Fig. 8.3: a) MSD analysis of the trajectories of particles crossing the ring canal. We have identified two groups of trajectories (green and orange) based on the shape of the MSD curve. b) Corresponding particles displayed in a cropped image of the ring canal. We show that the MSD groups of trajectories correspond to a specific spatial distribution of the particles : fast particles cross at the center of the canal, while slow particles cross near the canal rim. The motion differences comes from different mechanisms supporting the active transport of units.

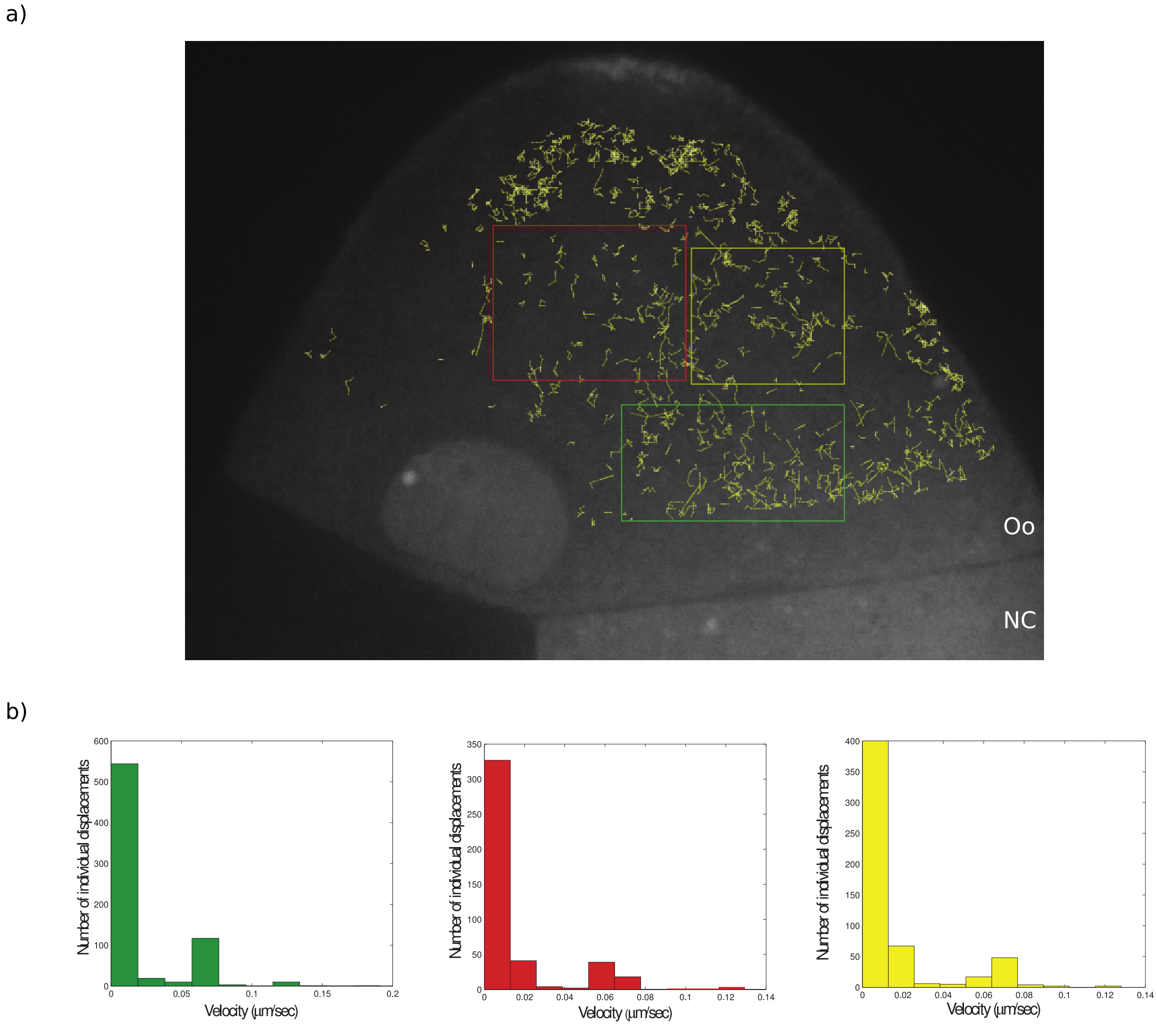


FIG. 8.4: 3D plus time tracking of *oOskar* mRNA. a) Final tracks superimposed on the final time point image. b) Histograms of velocity for the trajectories lying in three regions of the oocyte (green, red and yellow boxes in a)). The histograms show the bimodal distribution of the velocity of movements, whatever the region of the oocyte.

8.2 3D tracking of *Oskar* mRNA

The transport and localization of messenger RNAs (mRNAs) is essential in diverse processes such as the growth and the differentiation of eucaryotes such as the Drosophila. More specifically, the osk mRNA is known to play an important role in many nuclear and cytoplasmic events but its transport was poorly understood before a study to which we have participated [120]. This study, extensively involving imaging and particle tracking, has indeed allowed to refine the model of osk mRNA transport, answering questions such as the biophysical properties of the transport, the temporal requirement of specific nuclear or cytoplasmic proteins, and the principal nature of the osk during its transport.

As an example of experiments, molecular beacons targeting osk mRNA have been introduced in the ovocyte, 3D DSCM image sequences have been acquired, and we have tracked the mRNA complexes in the oocyte thanks to the E-MHT algorithm. We show in Figure 8.4 the resulting

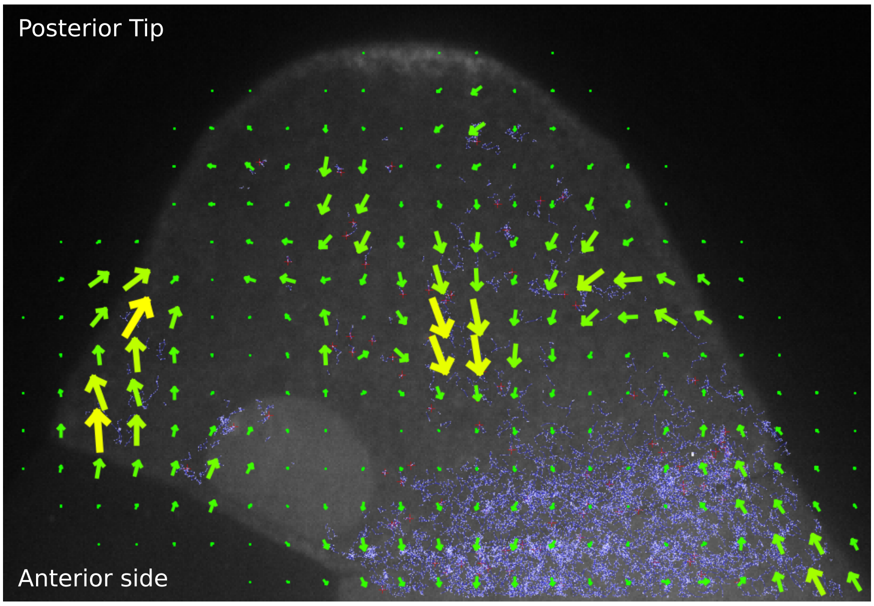


FIG. 8.5: Flow visualization of the Osk mRNA in the Drosophila oocyte. The flow is obtained by Gaussian smoothing of the measured movements. It shows a that near the border of the oocyte the osk particles move towards the posterior tip, with both flows then seeming to *rebound* from the posterior to the center creating a strong central anterior flow.

tracks superimposed on one slice of the final time point stack. Three regions of the oocyte were selected to analyze the velocity of osk and we show in Figure 8.4 the velocity histograms computed for each area. The histograms show that osk motion is heterogeneous with many molecules engaged in slow motion as well as a small sub-population engaged in rapid motion at a given time. Discriminating the two sub-populations would not have been feasible with a more global approach avoiding the study of single particles.

In Figure 8.5 we show the velocity and trend of flow of osk particles overlaid on the oocyte by arrows. The flow computation is achieved by a Gaussian smoothing of the measured particle displacements between subsequent frames, so large arrows represent faster velocities and stronger flows, whilst smaller arrows represent slower velocities. The flow visualization has helped to show that osk particles seem to have a posterior bias in the oocyte : osk particles concentrate at the posterior tip of the oocyte, with flows then seeming to *rebound* from the posterior to the center creating a strong central flow towards the anterior side of the oocyte.

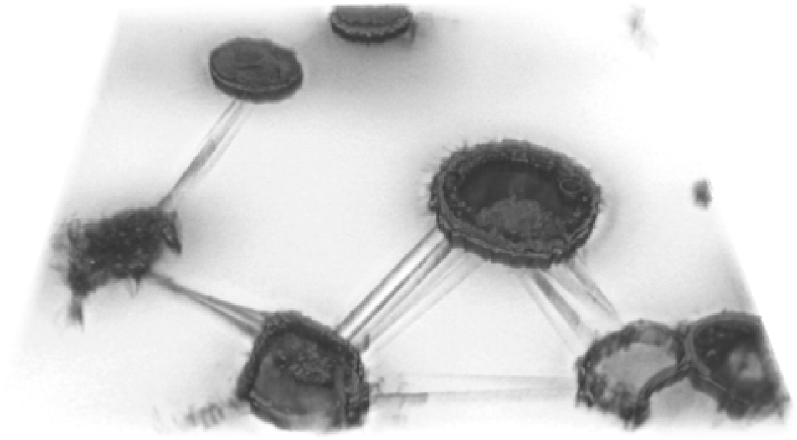


FIG. 8.6: Tunneling nanotubes linking neuronal cells in 3D fluorescence microscopy. CAD cells have been labelled with LysoTracker Red and observed with a DSCM.

8.3 Prions tracking inside tunneling nanotubes

In a recent study, to which we have collaborated, it was demonstrated that tunneling nanotubes (TNTs) allow transfer of exogenous and endogenous prions between infected and non-infected neuronal CAD cells [78]. The study has extensively exploited fluorescence microscopy techniques : the CAD cells were labeled with LysoTracker red and imaged with a DSCM in three dimensions. LysoTracker-positive vesicles were seen moving towards non-infected cells inside nanotubes, as shown in Figure 8.6. Characterizing prions motion on TNTs was wanted to understand their spreading mechanisms. However, automatically tracking prion complexes through time is complicated by the photons emitted by the TNTs. Indeed, TNTs appear as bright structures which make prions barely visible and discard particle detection techniques based on a flat background image model.

To facilitate prions identification we have applied the MCA source separation technique explained in Chapter 4 to the set of fluorescent images. We show in Figure 8.7 the separation of sources we have achieved.

In the particle images the infectious vesicles were automatically detected while moving inside TNTs and entering into the cells, and their trajectories were reconstructed thanks to E-MHT algorithm. A typical 3-D prion vesicle trajectory is shown in Figure 8.8. The MSD analysis of the track, also shown in Figure 8.8 in logarithmic scales. MSD data points of the LysoTracker vesicle in the TNT (red boxes) and in the cell (green triangles) are well fitted by a power law relationship with respect to the time intervals Δt, as shown by the good linear approximation when working in logarithmic scales (red and green solid lines). The slopes of these two lines are above 1, in the TNT (red, slope = 1.823) and in the cell (green, slope = 1.544), demonstrating that the vesicle has a directed movement throughout the movie even after entering the cell. The shift in the Y-axis between the red and green lines shows a slower directed movement after the vesicle enters the cell, which suggests the use of different transport mechanisms. Inside the TNT the estimated velocity of the tracked vesicle (41.5 nm/s on average) is very similar to that

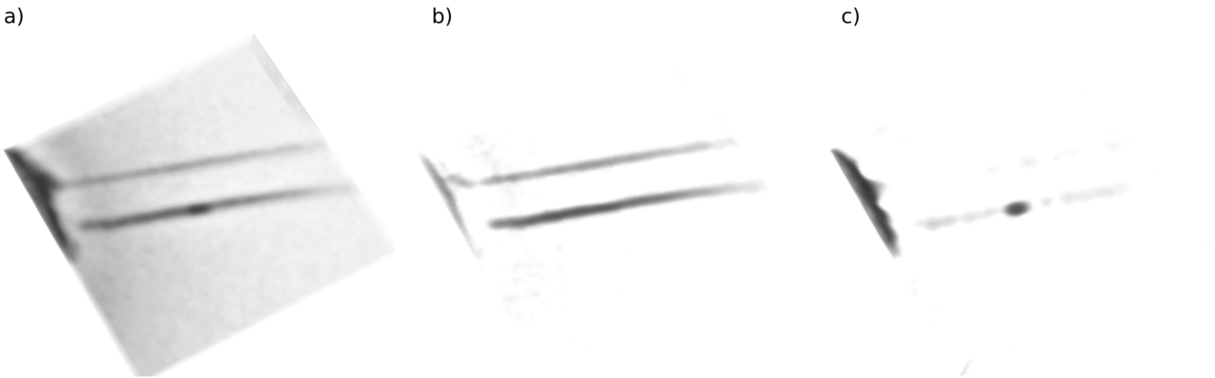


FIG. 8.7: Source separation of images of tunneling nanotubes linking neuronal cells. a) A cropped stack focusing on two TNTs and a prion vesicle. b) The image reconstructed from curvelet coefficients extracted by the MCA algorithm. c) The particle image extracted by the MCA algorithm.

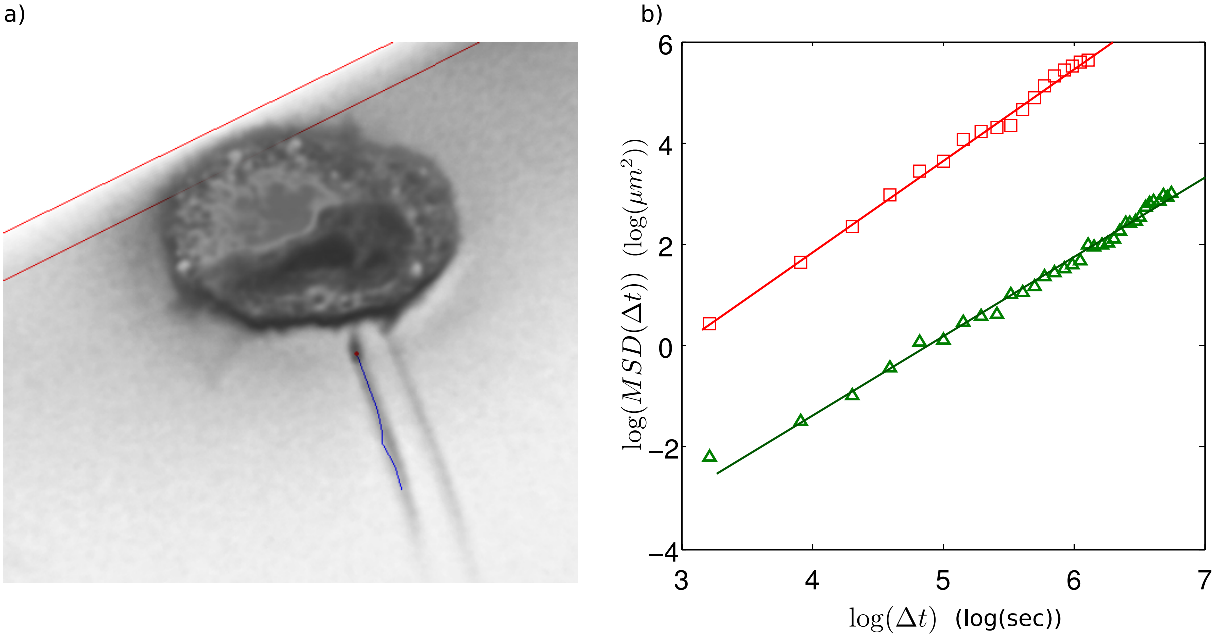


FIG. 8.8: a) Tracking of a prion (red sphere) inside a TNT before entering a recipient cell. b) MSD analysis of the vesicle's trajectory in logarithmic scales. We split the trajectory in two parts : moving inside the TNT (red boxes), and entering into the cell (green triangles). The red and green curves correspond to the linear approximations of the two sets of data points respectively.

calculated for a complex myosin VI and the GLUT1 transporter binding protein, indicating the possible involvement of actin-mediated motors.

Conclusions and perspectives

Particle tracking is the method of choice to characterize the dynamics of moving objects at the nanometric scale since it provides a robust and accurate means to understand sub-cellular mechanisms. Particle tracking in realistic biological conditions is a challenge to many image processing and analysis algorithms. In this work we have addressed several aspects related to the problem of tracking several hundreds of particles in cluttered conditions. In a first part we browsed through the existing solutions to the issue of tracking nanometric scale particles in biological images. This comprehensive study allowed us to identify some major limitations of standard solutions that hamper their use in a wide range of biological applications. The most prominent issue of advanced particle tracking techniques at the time this work was started is their lack of flexibility which specializes them to a reduced number of cases consistent with limitating assumptions. On the other hand, more general methods were unable to solve the most challenging issues of particle tracking in biology. The second part of this work has thus focused on designing new methodological solutions for particle tracking which achieve high quality results in a wealth of biological cases. We show in Figure 9 the two-step particle tracking procedure resulting from our contributions. As compared to the original procedure described in the introduction (Fig. 1), we show that both the detection and tracking steps have benefited from improvements, and that the standard procedure has been modified to introduce new interactions between the two steps. The last part of this work consisted in applying our new algorithms to a number of original and challenging biological applications.

We have first addressed the issue of detecting particles in fluorescence images containing a structured background. The key idea of the proposed method is to use the MCA algorithm [164] to separate the background from the particle signal by exploiting the differences between their morphologies in the image. We have made a number of adaptations to the MCA to comply with the characteristics of biological fluorescence images. For instance, we have proposed to use the curvelet dictionary and a wavelet dictionary with sparsity priors to split the background and the particles signals. After source separation, the background-free image can be reliably analyzed to identify the particle locations and track them through time. The proposed detection algorithm has found a direct application in a recent biological study investigating the spread of prions across cells by means of tunneling nanotubes. In these 3D fluorescence images the nanotubes intensity was obliterating the prion particle signal, hence making standard detection techniques inefficient. Using the the proposed procedure has allowed us to restore the particle signal and to robustly detect and track prions while entering into a cell.
This work on source separation in fluorescence images can be extended along several lines by further adapting the MCA to the characteristics of 3D+t fluorescence images. Using the MCA for processing large 3D images is still an open issue because of the difficulty to represent sparsely complex 3D backgrounds, and because of the computational cost of existing 3D transformations. These issues have led us to process the stacks slice-by-slice. Yet, with efforts being conducted to

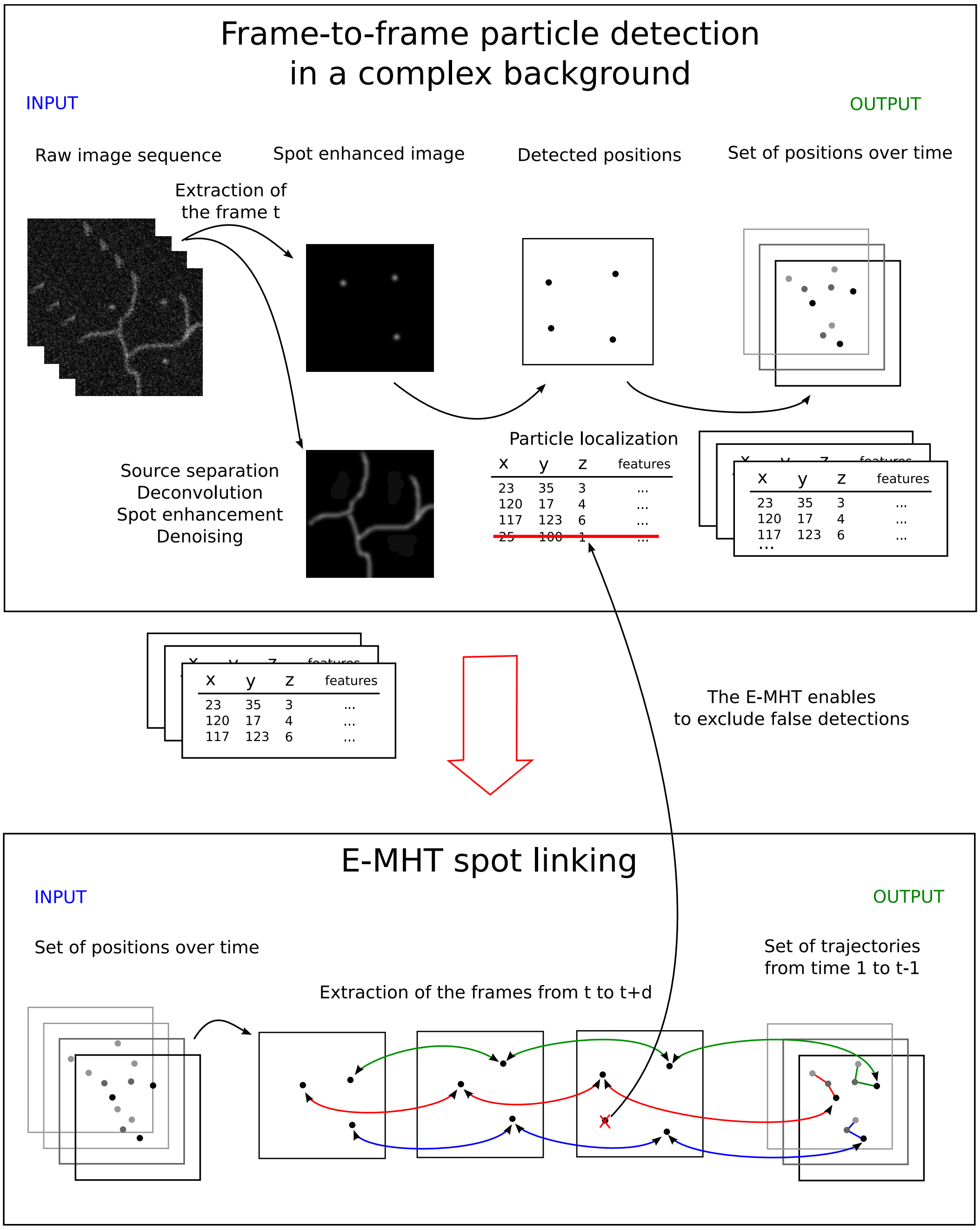


FIG. 9: Novel particle tracking procedure including the MCA-UDWT for detecting particles and the E-MHT for linking detections through time. The E-MHT is able to correct the false detections outputted by the detection step by automatically discarding detections which are not consistent with past and future frames.

improve 3D representation dictionaries [115, 184], we shall apply the proposed approach in a full 3D way. Accounting for the 3D morphologies of objects would result in a substantially improved source separation.

The second problem we have discussed is the modeling of the multiple particle tracking problem in cluttered conditions. We argue that a global modeling is required to achieve a high

degree of interpretability of the methods, while standard techniques are often very local, with scores and parameters which are difficult to interprete in a global sense. We have thus proposed a global statistical framework which accounts for every aspect of the particle tracking problem. The designed probabilistic framework includes several models which are dedicated to biological imaging, such as statistical models of motion of particles. We have also defined the concept of target perceivability [110] for biological particles. By doing so, the existence of particles is explicitly modeled and statistically quantified, thereby addressing the issues of track termination/creation within the statistical framework for tracking. As a result, the proposed framework enjoys a high degree of flexibility since every parameter and model involved finds a simple and intuitive interpretation. The proposed probabilistic framework has consequently allowed us to exhaustively model a wealth of different biological cases.
Accounting for more models dedicated to different biological cases within the probabilistic framework would widen further its range of applications. For instance, the perceivability model can be extended to explicitly take into account unusual intensity profiles, such as blinking particles which are recurrently missed by the detection procedure. Moreover, in the current solution we only use the location of detections in images as an input of the tracking method, but in some cases richer information can be exploited. As an example, a preliminary work has been dedicated to accounting for the the target intensity profiles when tracking [39, 40]. It has allowed us separating closely spaced particles which give birth to a single detection. Extending further this approach and integrating it in the proposed probabilitic framework would result in important improvements.

Our third contribution is the reformulation of the MHT algorithm to include the probalistic framework for tracking biological particles, and the design of an efficient implementation which allows one to track numerous particles in poor imaging conditions. The probabilistic framework allows us to explicitly model within the algorithm various random events, such as target appearance/disappearance and the corruption of the detection sets. The E-MHT takes full advantage of the tracking model by incorporating the knowledge from future frames, whereby the significance of statistical scores is increased and the reliability of the tracking is improved. As an example, the E-MHT is able to automatically discard false detections by analyzing their consistency with detections in past and future frames. The increased complexity resulting from the use of future frames is addressed by an efficient design of the algorithm which exploits the tree organization of the solutions and the ability to make the computations in a parallel manner. A series of comparative tests between the E-MHT and a number of state-of-the-art techniques have been performed with synthetic 2D image sequences and real 2D and 3D data sets. In every case the E-MHT has proved superior performance over standard techniques, with a remarkable capability to handle very poor imaging conditions thanks to its ability to discriminate true detections from false ones. The ability of the E-MHT to track numerous particles despite the poor imaging conditions has been illustrated in two different challenging biological applications : the tracking of *Oskar* mRNA in 3D and the trafficking of Golgi units, both in the Drosophila oocyte.
As shown in Figure 9, the E-MHT algorithm introduces novel interactions between the detection and the tracking steps since it is able to correct the output of the detection procedure. A future development of our approach is to modify further the standard two-step tracking procedure to enforce the two-ways interactions between the image and the tracking models. As an example, allowing the tracking algorithm to suggest some locations to the detection procedure is a promising development. A challenging continuation of this work would be to fuse in a common probabilistic framework the two steps constituing the standard approach for tracking, while maintaining the high degree of flexibility achieved by the E-MHT.

Part IV

Annexes

Annex A

Experimental setups

A.1 Computer setup

We have mainly used three setups for the different experiences presented in the manuscript and we detail them next.
Setup 1 and Setup 2 have multicore cpus and large memory capacities which makes them suited to large datasets processing with computing power demanding algorithms. They also prefigure the standard computer setups in most labs in few years : the roadmap of computer processor constructors clearly indicates that the multicore cpus will widespread and that the number of cores will still grow. The memory capacities are also considerably growing.
The setup 3 is a more modest configuration, which is useful for testing the feasibility of algorithms on computers found in most biological labs.

Setup 1 Sun Ultra 40 M2 Workstation with Debian 5.0.3 kernel Linux 2.6.26-2-amd64.
Processor : two dual-core AMD Opteron 2000 series (=4 cores).
Memory : 12GB of 667MHz DDR2 ECC SDRAM.

- Matlab®R2008b from the MathWorks ™.
- Java™SE Runtime Environment (build 1.6.0_14)

Setup 2 Apple Mac Pro Quad with Mac OS X 10.6.1.
Processor : two 2.26 GHz (E5520) Quad-core Intel Xeon "Gainestown" (= 8 cores).
Memory : 10GB of 1066MHz DDR3 ECC SDRAM.
Environment :

- Matlab®R2007a from the MathWorks ™.
- Java™SE Runtime Environment (build 1.6.0_15)

Setup 3 Apple MacBook pro with Mac OS X 10.6.1.
Processor : 2.4 GHz Intel Core 2 Duo (=2 cores).
Memory : 4GB 667 MHz DDR2 SDRAM.
Environment :

- Matlab®R2007a from the MathWorks ™.
- Java™SE Runtime Environment (build 1.6.0_15)

Annex B

Additional results

B.1 Enumerating feasible associations for a simple multiple target example

Enumeration of all the feasible solutions of the association problem shown in Figure 5.1 :

$$
\begin{aligned}
\Omega^1 = \{ & \\
& \emptyset, \\
& \theta_1 : \{z_1(1)\}, \\
& \theta_1 : \{z_2(1)\}, \\
& \theta_1 : \{z_1(1)\} \cup \theta_2 : \{z_2(1)\} \\
& \},
\end{aligned}
$$

for only one frame. For two frames the set of feasible associations contains 23 solutions :

$$
\begin{aligned}
\Omega^2 = \Omega^1 \cup \{ & \\
& \theta_1 : \{z_1(1)\} \cup \theta_2 : \{z_1(2)\}, \\
& \theta_1 : \{z_1(1)\} \cup \theta_2 : \{z_2(2)\}, \\
& \theta_1 : \{z_1(1)\} \cup \theta_2 : \{z_2(1), z_1(2)\}, \\
& \theta_1 : \{z_1(1)\} \cup \theta_2 : \{z_2(1), z_2(2)\}, \\
& \theta_1 : \{z_1(1), z_1(2)\}, \\
& \theta_1 : \{z_1(1), z_1(2)\} \cup \theta_2 : \{z_2(1)\}, \\
& \theta_1 : \{z_1(1), z_1(2)\} \cup \theta_2 : \{z_2(1), z_2(2)\}, \\
& \theta_1 : \{z_1(1), z_2(2)\}, \\
& \theta_1 : \{z_1(1), z_2(2)\} \cup \theta_2 : \{z_2(1)\}, \\
& \theta_1 : \{z_1(1), z_2(2)\} \cup \theta_2 : \{z_2(1), z_1(2)\}, \\
& \theta_1 : \{z_2(1)\} \cup \theta_2 : \{z_1(2)\}, \\
& \theta_1 : \{z_1(1)\} \cup \theta_2 : \{z_2(2)\}, \\
& \theta_1 : \{z_2(1), z_1(2)\}, \\
& \theta_1 : \{z_2(1), z_1(2)\} \cup \theta_2 : \{z_2(2)\}, \\
& \theta_1 : \{z_2(1), z_2(2)\}, \\
& \theta_1 : \{z_2(1), z_2(2)\} \cup \theta_2 : \{z_1(2)\}, \\
& \theta_1 : \{z_1(2)\}, \\
& \theta_1 : \{z_2(2)\}, \\
& \theta_1 : \{z_1(2)\} \cup \theta_2 : \{z_2(2)\} \\
& \}.
\end{aligned}
$$

B.2 Probabilistic distribution of minimal distances between detections

We aim here at defining a statistical model for distances between detections that do not originate from the same target. We assume that p detections, forming a set $Z = \{\mathbf{z}_i\}_{i=1..p}$, are randomly localized in an image I following an uniform spatial distribution. Given another detection $\mathbf{z}$ in I, we study the pdf of the Euclidian distance, d, between $\mathbf{z}$ and its closest detection in the set Z :

$$d \triangleq \min_{i=1..p} ||\mathbf{z} - \mathbf{z}_i||_2. \tag{B.1}$$

We can rewrite the event of observing d as the event of measuring one detection in Z at a distance d to $\mathbf{z}$, and finding every other detection at a greater distance as follows :

$$\exists i \in [1, p] \ s.t. \ ||\mathbf{z} - \mathbf{z}_i|| = d \text{ and } \forall j = 1 \dots p : ||\mathbf{z} - \mathbf{z}_j|| \geq d. \tag{B.2}$$

We use this last formulation to define a pdf for d.
We assume first, without loss of generality, that the detection $\mathbf{z}_1$ is at a distance d to $\mathbf{z}$. Since $\mathbf{z}_1$ is uniformly distributed in I, the pdf of $\mathbf{z}_1$ being at a distance d from $\mathbf{z}$ is the ratio of the volume of the subdomain of I which is at a distance d from $\mathbf{z}$, to the volume of I : V. In 3D the

ensemble of points at a distance d to $\mathbf{z}$ is a sphere centered in $\mathbf{z}$ of radius d : $\mathcal{S}_d$. So, the pdf of the distance d between $\mathbf{z}$ and $\mathbf{z}_1$ is :

$$f_1(d) = \frac{4\pi d^2}{V}, \tag{B.3}$$

if $\mathcal{S}_d$ is fully contained in I. In the following we neglect border effects : we assume that detections are distant from more than d to the image border, hence $\mathbf{S}_d$ is contained in I and the pdf defined in Equation B.3 is suited. This simplification is valid in most cases since the minimal distance between detections, d, is much smaller than the dimensions of I.

The detection $\mathbf{z}_1$ is the closest to $\mathbf{z}$ only if every other detections in Z are at a greater distance. The probability that one detection $\mathbf{z}_j$ with $1 < j \leq p$ is at a distance to $\mathbf{z}$ superior to d corresponds to the probability that it does not fall inside a ball which is centered in $\mathbf{z}$ with a radius d : $\mathcal{B}_d$. We thus write :

$$P\{||\mathbf{z} - \mathbf{z}_j|| > d\} = P\{\mathbf{z}_j \notin \mathcal{B}_d\} = 1 - \frac{4\pi d^3}{3V}. \tag{B.4}$$

The probability that the $p-1$ detections in $Z \setminus \{\mathbf{z_1}\}$ are all at a distance from $\mathbf{z}$ which is superior to d can therefore be written as :

$$P\{\forall j = 2..p,\ ||\mathbf{z} - \mathbf{z}_j|| > d\} = \Big(1 - \frac{4\pi d^3}{3V}\Big)^{p-1}. \tag{B.5}$$

The same calculations can be extended from the detection $\mathbf{z}_1$ to the p detections in Z in such a way that the pdf of the mininal distance between detections in Z and $\mathbf{z}$ is written as :

$$f(d) = p\frac{4\pi d^2}{V}\Big(1 - \frac{4\pi d^3}{3V}\Big)^{p-1}, \tag{B.6}$$

by combining Equations B.3 and B.5.

The Equation B.6 highlights an important benefit of studying only minimal distances between detections and not the distribution of every distances : the pdf we define depends only on the number of detections and the volume V, while accounting for the image shape is not required. In the case of minimal distances we indeed model only small neighborhoods of radius d around detections, while computing every distances would require describing the complete topology of the environment.

The demonstration is applicable in a straightforward way to 2D images. In this case, the pdf of the distance between a detection $\mathbf{z}_1$ and $\mathbf{z}$ is the ratio of the perimeter of the circle of radius r to V :

$$f_1(d) = \frac{2\pi d}{V}. \tag{B.7}$$

The probability that the $p-1$ detections in $Z \setminus \{\mathbf{z_1}\}$ are all at a distance from $\mathbf{z}$ which is superior to d corresponds to the probability that they all fall outside a disk of radius d :

$$P\{\forall j = 2..p,\ ||\mathbf{z} - \mathbf{z}_j|| > d\} = \Big(1 - \frac{\pi d^2}{V}\Big)^{p-1}. \tag{B.8}$$

By combining Equations B.7 and B.8 we thus write the pdf of the mininal distance between $\mathbf{z}$ and detections in Z as :

$$f(d) = p\frac{2\pi d}{V}\Big(1 - \frac{\pi d^2}{V}\Big)^{p-1}. \tag{B.9}$$

Publications

Journal publication

- Nicolas Chenouard*, Alexandre Dufour* and Jean-Christophe Olivo-Marin, "Tracking algorithms chase down pathogens", *Biotechnology Journal*, 4(6) :838-845, June 2009.
- Musa M. Mhlanga, Diana Bratu, Auguste Genovesio, Agata Rybarska, Nicolas Chenouard, Ulf Nehrbass, Jean-Christophe Olivo-Marin, "In vivo colocalisation of oskar mRNA and trans-acting proteins revealed by quantitative imaging of the drosophila oocyte", *PLoS ONE*, 4(7) :e6241, July 2009.
- Karine Gousset, Edwin Schiff, Christelle Langevin, Zrinka Marijanovic, Anna Caputo, Duncan Browman, Nicolas Chenouard, Fabrice de Chaumont, Angelo Martino, Jost Enninga, Jean-Christophe Olivo-Marin, Daniela Mannel and Chiara Zurzolo. "Prions hijack tunnelling nanotubes for intercellular spread.", Nature Cell Biology, 11 :328-336, February 2009.
- Emmanuelle Nicolas, Nicolas Chenouard, Jean-Christophe Olivo-Marin and Antoine Guichet, "A Dual Role for Actin and Microtubule Cytoskeleton in the Transport of Golgi Units from the Nurse Cells to the Oocyte Across Ring Canals", *Molecular Biology of the Cell*, 20(1) :556-568, January 2009.

Conference publication

- Nicolas Chenouard, Isabelle Bloch and Jean-Christophe Olivo-Marin, "Multiple Hypothesis Tracking in cluttered condition". *16th IEEE International Conference on Image Processing, ICIP 2009*, Cairo, Egypt, 2009.
- Nicolas Chenouard, Isabelle Bloch and Jean-Christophe Olivo-Marin, "Particle tracking in fluorescent microscopy images improved by morphological source separation". *16th IEEE International Conference on Image Processing, ICIP 2009*, Cairo, Egypt, 2009.
- Nicolas Chenouard, Isabelle Bloch and Jean-Christophe Olivo-Marin, "Multiple Hypothesis Tracking in microscopy images". *6th IEEE International Symposium on BioImaging, ISBI 2009*, Boston, USA, 2009.
- Nicolas Chenouard, Samantha Vernhettes, Isabelle Bloch and Jean-Christophe Olivo-Marin, "Morphological source separation for particle tracking in complex biological environments". *19th IAPR International Conference on Pattern Recognition, ICPR 2008*, Tampa, USA, 2008.
- Clovis Tauber, Pedro F. G. Rodriguez, Vincent Loriette, Nicolas Chenouard, Benoit Dubertret and Jean-Christophe Olivo-Marin, "Fast and automatic reconstruction of structured illumination microscopy images with multiscale products". *19th International Conference on Pattern Recognition, ICPR 2008*, Tampa, USA, 2008.

- Nicolas Chenouard, Fabrice de Chaumont, Isabelle Bloch and Jean-Christophe Olivo-Marin. "Improving 3D tracking in microscopy by joint estimation of kinetic and image models". *Proceedings of the international conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI 2008, MIAAB workshop*, New-York, USA, 2008.
- Nicolas Chenouard, Isabelle Bloch and Jean-Christophe Olivo-Marin, "Feature-aided particle tracking". *15th IEEE International Conference on Image Processing, ICIP 2008*, San Diego, USA, 2008.
- Fabrice de Chaumont, Nicolas Chenouard, Aurélie Mouret, Pierre Marie Lledo, Jean-Christophe Olivo-Marin, "A multi-threaded program architecture for an asynchronous and highly responsive GUI for automatic neuronal survival quantification", *5th IEEE International Symposium on Biomedical Imaging, ISBI 2008*, Paris, France, 2008.
- Bo Zhang, Nicolas Chenouard, Jean-Christophe Olivo-Marin and Vannary Meas-Yedid, "Statistical colocalization in biological imaging with false discovery control", *5th IEEE International Symposium on Biomedical Imaging, ISBI 2008*, Paris, France, 2008.

Bibliography

[1] P. Abrial, Y. Moudden, J.-L. Starck, J. Bobin, M.J. Fadili, B. Afeyan, and M. Nguyen. Morphological component analysis and inpainting on the sphere : Application in physics and astrophysics. *Journal of Fourier Analysis and Applications (JFAA), special issue on "Analysis on the Sphere"*, 13(6) :729–748, December 2007.

[2] F. Aguet, D. Van de Ville, and M. Unser. A maximum-likelihood formalism for sub-resolution axial localization of fluorescent nanoparticles. *Optics Express*, 13 :10503–10522, December 2005.

[3] R. Amino, A. Genovesio, S. Celli, P. Roux, J.-C. Olivo-Marin, S. L. Shorte, R Menard, and F. Frischknecht. From skin to liver : in vivo imaging of plasmodium sporozoites. In *Molecular Parasitology Meeting XV*, September 19-23 2004.

[4] C.M. Anderson, G.N. Georgiou, I.E. Morrison, G.V. Stevenson, and R.J. Cherry. Tracking of cell surface receptors by fluorescence digital imaging microscopy using a charge-coupled device camera. low-density lipoprotein and influenza virus receptor mobility at 4 degrees c. *Journal of Cell Science*, 101(2) :415–425, February 1992.

[5] J. Apgar, Y. Tseng, E. Federov, M. Herwig, S. Almo, and D. Wirtz. Multiple-particle tracking measurements of heterogeneities in solutions of actin filaments and actin bundles. *Biophysical Journal*, 79 :1095–1106, 2000.

[6] N. Arhel, A. Genovesio, K.A. Kim, S. Miko, E. Perret, J.-C. Olivo-Marin, S. Shorte, and P. Charneau. Quantitative four-dimensional tracking of cytoplasmic and nuclear HIV-1 complexes. *Nature Methods*, 3 :817–24, 2006.

[7] M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2) :174–188, 2002.

[8] R.H. Bamberger and M.J.T. Smith. A filter bank for the directional decomposition of images : theory and design. *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, 40(4) :882–893, Apr 1992.

[9] Y. Bar-Shalom. Dimensionless score function for multiple hypothesis tracking. *Aerospace and Electronic Systems, IEEE Transactions on*, 43(1) :392–400, January 2007.

[10] Y. Bar-Shalom, S. S. Blackman, and R. J. Fitzgerald. The dimensionless score function for multiple hypothesis decision in tracking. In *Systems, Man and Cybernetics, 2005 IEEE International Conference on*, volume 4, pages 3264–3269 Vol. 4, Oct. 2005.

[11] Y. Bar-Shalom and W. Dale Blair. *Multitarget-multisensor tracking applications and advances*, volume III. Artech House, Oct 2000.

[12] Y. Bar-Shalom, K.C. Chang, and H.A.P. Blom. Tracking a maneuvering target using input estimation versus the interacting multiple model algorithm. *Aerospace and Electronic Systems, IEEE Transactions on*, 25(2) :296–300, Mar 1989.

[13] Y. Bar-Shalom and T.E. Fortmann. *Tracking and data association.* Academic Press, New-York, 1988.

[14] Y. Bar-shalom, Rong X. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation : Theory Algorithms and Software (A Wiley-Interscience publication).* Wiley-Blackwell, July 2001.

[15] J. E. Bear, T. M. Svitkina, M. Krause, D. A. Schafer, J. J. Loureiro, G. A. Strasser, I. V. Maly, O. Y Chaga, J. A. Cooper, G. G. Borisy, and F. B. Gertle. Antagonism between ena/vasp proteins and actin filament capping regulates fibroblast motility. *Cell*, 109(4) :509–521, 2002.

[16] Y. Benjamini and D. Yekutieli. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 29 :1165–1188, 2001.

[17] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess. Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science*, 313(5793) :1642–1645, 2006.

[18] S. S. Blackman. Multiple hypothesis tracking for multiple target tracking. *Aerospace and Electronic Systems Magazine, IEEE*, 19(1) :5–18, 2004.

[19] S. S. Blackman, R. J. Dempster, and R. W. Reed. Demonstration of multiple-hypothesis tracking (mht) practical real-time implementation feasibility. In Oliver E. Drummond, editor, *Signal and Data Processing of Small Targets 2001*, volume 4473, pages 470–475. SPIE, 2001.

[20] S. S. Blackman and R. Popoli. *Design and Analysis of Modern Tracking Systems.* Artech House, Norwood, MA, USA, 1999.

[21] H. A. P. Blom and Y. Bar-Shalom. The interacting multiple model algorithm for systems with markovian switching coefficients. *IEEE Transactions on Automatic Control*, 33(8) :780–783, August 1988.

[22] J. Bobin, Y. Moudden, J.-L. Starck, and M. Elad. Morphological diversity and source separation. *Signal Processing Letters, IEEE*, 13(7) :409–412, July 2006.

[23] J. Bobin, J.-L. Starck, J. Fadili, and Y. Moudden. Sparsity and morphological diversity in blind source separation. *Image Processing, IEEE Transactions on*, 16(11) :2662–2674, november 2007.

[24] J. Bobin, J.-L. Starck, Y. Moudden, and J. Fadili. Blind source separation : The sparsity revolution. *Advances in Imaging and Electron Physics*, 152 :221–306, 2008.

[25] S. Bonneau, M. Dahan, and L. D. Cohen. Single quantum dot tracking based on perceptual grouping using minimal paths in a spatiotemporal volume. *IEEE Transactions on Image Processing*, 14(9) :1384–1395, September 2005.

[26] C. Bouzigues and M. Dahan. Transient directed motions of gabaa receptors in growth cones detected by a speed correlation index. *Biophysical Journal*, 92(2) :654 – 660, january 2007.

[27] A. W. Bowman and A. Azzalini. *Applied Smoothing Techniques for Data Analysis : The Kernel Approach with S-Plus Illustrations.* Oxford Statistical Science Series. Oxford University Press, USA, November 1997.

[28] B. Brandenburg and X. Zhuang. Virus trafficking - learning from single-virus tracking. *Nature Reviews Microbiology*, 5(3) :197–208, March 2007.

[29] X. Breen, G.H. Joss, and K.L. Williams. Locating objects of interest within biological objects : the top-hat box filter. *Comp. Ass. Microsc.*, 2(3) :97–102, 1991.

[30] D. S. Bright and E. B. Steel. Two-dimensional top hat filter for extracting spots and spheres from digital images. *Journal of Microscopy*, 146 :191–200, 1987.

[31] R. E. Burkard. Selected topics on assignment problems. *Discrete Appl. Math.*, 123(1-3) :257–302, 2002.

[32] P. Burt and E. Adelson. The laplacian pyramid as a compact image code. *Communications, IEEE Transactions on*, 31(4) :532–540, Apr 1983.

[33] G. G. Cabal, A Genovesio, S. Rodriguez-Navarro, C. Zimmer, O. Gadal, A. Lesne, H. Buc, F. Feuerbach-Fournier, J.-C. Olivo-Marin, E.C. Hurt, and U. Nehrbass. SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope. *Nature*, 441(7094) :770–773, June 2006.

[34] E. Candes and D. Donoho. Curvelets - a surprisingly effective nonadaptive representation of objects with edges. Technical report, California Institute of Technology, 1999.

[35] D. L. Candes E. J., Donoho. Wavelets : The key to intermittent information ? *Philosophical Transactions : Mathematical, Physical and Engineering Sciences*, 357 :2495–2509, September 1999.

[36] J.-F. Cardoso. Blind signal separation : statistical principles. *Proceedings of the IEEE*, 86(10) :2009–2025, 1998.

[37] M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher. Green fluorescent protein as a marker for gene expression. *Science*, 263(5148) :802–805, 1994.

[38] Michael K. Cheezum, William F. Walker, and William H. Guilford. Quantitative comparison of algorithms for tracking single fluorescent particles. *Biophysical Journal*, 81 :2378–2388, October 2001.

[39] N. Chenouard, N. Bloch, and J.-C. Olivo-Marin. Feature-aided particle tracking. In *Proceedings of the IEEE International Conference on Image Processing*, San Diego, October 2008. IEEE.

[40] N. Chenouard, F. de Chaumont, I. Bloch, and J.-C. Olivo-Marin. Improving 3d tracking in microscopy by joint estimation of kinetic and image models. In *Proceedings of the international conference on Medical Image Computing and Computer-Assisted Intervention - MICCAI, , MIAAB workshop*, New York, NY, USA, October 2008.

[41] N. Chenouard, A. Dufour, and J.-C. Olivo-Marin. Tracking algorithms chase down pathogens. *Biotechnology Journal*, 4(6) :838–845, 2009.

[42] D. Chetverikov and J. Verestoy. Feature point tracking for incomplete trajectories. *Computing*, 62(4) :321–338, 1999.

[43] A. Cichocki and S. Amari. *Adaptive Blind Signal and Image Processing : Learning Algorithms and Applications.* John Wiley & Sons, Inc., New York, NY, USA, 2002.

[44] L. D. Cohen and R. Kimmel. Global minimum for active contour models : A minimal path approach. *Int. J. Comput. Vision*, 24(1) :57–78, 1997.

[45] D. Comaniciu and P. Meer. Mean shift : a robust approach toward feature space analysis. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 24, pages 603–619, May 2002.

[46] P. Comon. Independent component analysis, a new concept ? *Signal Process.*, 36(3) :287–314, 1994.

[47] I. J. Cox and S. L. Hingorani. An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(2) :138–150, Feb 1996.

[48] Ingemar J. Cox. A review of statistical data association for motion correspondence. *Int. J. Comput. Vision*, 10(1) :53–66, 1993.

[49] J. C. Crocker and D. G. Grier. Methods of digital video microscopy for colloidal studies. *Journal of Colloid and Interface Science*, 179 :298–310, 1996.

[50] S. Deb, M. Yeddanapudi, K. Pattipati, and Y. Bar-Shalom. A generalized s-d assignment algorithm for multisensor-multitarget state estimation. *Aerospace and Electronic Systems, IEEE Transactions on*, 33(2) :523–538, April 1997.

[51] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1) :1–38, 1977.

[52] M.N. Do and M. Vetterli. Pyramidal directional filter banks and curvelets. In *Image Processing, 2001. Proceedings. 2001 International Conference on*, volume 3, pages 158–161 vol.3, 2001.

[53] D. L. Donoho and M. R. Duncan. Digital curvelet transform : Strategy, implementation and experiments. In *in Proceedings of Aerosense 2000, Wavelet Applications VII*, pages 12–29. SPIE, 1999.

[54] A. Doucet, N. De Freitas, and N. J. Gordon, editors. *Sequential Monte Carlo methods in practice.* Springer-Verlag, 2001.

[55] A. Doucet, S. Godsill, and C. Andrieu. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and computing*, 10(3) :197–208, 2000.

[56] B. Dubertret, P. Skourides, D. J. Norris, V. Noireaux, A. H. Brivanlou, and A. Libchaber. In Vivo Imaging of Quantum Dots Encapsulated in Phospholipid Micelles. *Science*, 298(5599) :1759–1762, 2002.

[57] D. L. Donoho E. J. Candes, L. Demanet and L. Ying. Fast discrete curvelet transforms. *SIAM Journal on Multiscale Modeling and Simulation*, 3 :861 – 899, 2006.

[58] E. L. Elson. Fluorescence correlation spectroscopy and photobleaching recovery. *Annual Review of Physical Chemistry*, 36(1) :379–406, 1985.

[59] H. Eslami, R. ; Radha. Translation-invariant contourlet transform and its application to image denoising. *Image Processing, IEEE Transactions on*, 15(11) :3362–3374, November 2006.

[60] M.J. Fadili and J.-L. Starck. Em algorithm for sparse representation-based image inpainting. In *IEEE International Conference on Image Processing, 2005. ICIP 2005.*, volume 2, pages II–61–4, September 2005.

[61] M.J. Fadili and J.-L. Starck. Image decomposition and separation using sparse representations : an overview. In *Proceedings of the IEEE Special Issue on : Applications of Sparse Representation and Compressive Sensing*, 2010. in press.

[62] Y. Freund and R. E. Schapire. A short introduction to boosting. *Journal of the Japanese Society for Artificial Intelligence*, 14(5) :771–780, September 1999.

[63] A. Genovesio. Multitarget tracking applications in microbiology. In *ELSO's annual congress*, Nice (France), September 4-8 2004.

[64] A. Genovesio. *Une méthode de poursuite de tâches multiples Application à l'étude de la dynamique d'objets biologiques en microscopie 3D+T*. PhD thesis, Universit'e Paris 5, 2005.

[65] A. Genovesio, Z. Belhassine, and J.-C. Olivo-Marin. Adaptive Gating in Gaussian bayesian Multi-Target Tracking. In *Proceedings of the IEEE International Conference on Image Processing*, volume 1, pages 147–150, Octobre 2004.

[66] A. Genovesio, G. Cabal, O. Gadal, F. Feuerbach, U. Nehrbass, and J.-C. Olivo-Marin. A Computerize Method for the Three-dimensional detection and Localization of Chromosomal Loci in Living Cells. In *Proceedings of the European Light Microscopy Initiative meeting*, Gothenburg, Sweden, 26-28 May 2004.

[67] A. Genovesio, T. Liedl, V. Emiliani, W. J. Parak, M. Coppey-Moisan, and J.-C. Olivo-Marin. Multiple particle tracking in 3D+T microscopy : methods and application to the tracking of endocytosed quantum dots. *IEEE Transactions on Image Processing*, 15(5) :1062–1070, 2006.

[68] A. Genovesio and J.-C. Olivo Marin. Interacting Multiple Model Based Method to Track Moving Fluorescent Biological Spots. In *Proceedings of the International Symposium on Biomedical Imaging*, pages 1239–1242, 2004. Arlington (USA).

[69] A. Genovesio and J.-C. Olivo Marin. Split and Merge data Association Filter for Dense Multi-Target tracking. In *Proceedings of the International Conference on Pattern recognition*, August 2004. Cambridge (UK).

[70] A. Genovesio and J.-C. Olivo-Marin. Tracking fluorescent spots in biological video microscopy. In J.-A. Conchello, C.J. Cogswell, and T. Wilson, editors, *Proceedings of SPIE - Three-Dimensional and Multidimensional Microscopy : Image Acquisition and Processing X*, volume 4964, pages 98–105, January 2003. San Jose, California, USA.

[71] A. Genovesio, B. Zhang, and J.-C. Olivo-Marin. Tracking of Multiple Fluorescent Biological Objects in Three Dimensional Video Microscopy. In *Proceedings of the IEEE International Conference on Image Processing*, volume I, pages 1105–1108, September 2003. Barcelona, Spain.

[72] A. Genovesio, Bo Zhang, Vannary Meas-Yedid, Christophe Zimmer, and Jean-Christophe Olivo-Marin. Tracking and motion analysis in dynamic imaging : movie crunching. In *Proceedings of the European Light Microscopy Initiative meeting*, June 11-13 2003. Barcelone (Spain).

[73] W. R. Gilks and C. Berzuini. Following a moving target – Monte Carlo inference for dynamic Bayesian models. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, pages 127–146, 2001.

[74] W. J. Godinez, M. Lampe, S. Worz, B. Muller, R. Eils, and K. Rohr. Probabilistic tracking of virus particles in fluorescence microscopy images. *Biomedical Imaging : From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pages 272–275, May 2008.

[75] W.J. Godinez, M. Lampe, S. Worz, B. Muller, R. Eils, and K. Rohr. Tracking of virus particles in time-lapse fluorescence microscopy image sequences. In *Biomedical Imaging : From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on*, pages 256–259, April 2007.

[76] N. Gordon. A hybrid bootstrap filter for target tracking in clutter. In *American Control Conference, 1995. Proceedings of the*, volume 1, pages 628–632 vol.1, Jun 1995.

[77] N. J. Gordon. Non-linear/non-gaussian filtering and the bootstrap filter. In *Non-Linear Filters, IEE Colloquium on*, pages 4/1–4/6, May 1994.

[78] K. Gousset, E. Schiff, C. Langevin, Z. Marijanovic, A. Caputo, D. T. Browman, N. Chenouard, F. de Chaumont, A. Martino, J. Enninga, J.-C. Olivo-Marin, D. Männel, and C. Zurzolo. Prions hijack tunneling nanotubes for intercellular spread. *Nature cell biology*, 11 :328–336, February 2009.

[79] D. A. Grossman and O. Frieder. *Information Retrieval : Algorithms and Heuristics.* The Kluwer International Series of Information Retrieval. Springer, P.O.Box 17, 3300 AA Dordrecht, The Netherlands, zweite edition, 2004.

[80] M. Gu. *Advanced optical imaging theory.* Springer-Verlag, Berlin, 2000.

[81] S. L. Gupton, W. C. Salmon, and C. M. Waterman-Storer. Converging populations of f-actin promote breakage of associated microtubules to spatially regulate microtubule turnover in migrating cells. *current biology*, 12(22) :1891–1899, 2002.

[82] M.G Gustafsson. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of Microscopy*, 198 :82–87, 2000.

[83] S. W. Hell. Toward fluorescence nanoscopy. *Nature Biotechnology*, 21(11) :1347–1355, November 2003.

[84] F. L. Hitchcock. The distribution of a product from several sources to numerous localities. *Journal of Math. Phys.*, 20 :224–230, 1941.

[85] M. Holschneider, R. Kronland-Martinet, J. Morlet, and P. Tchamitchian. A Real-Time Algorithm for Signal Analysis with the Help of the Wavelet Transform. In J.-M. Combes, A. Grossmann and P. Tchamitchian, editor, *Wavelets. Time-Frequency Methods and Phase Space*, pages 286–297. Springer-Verlag, 1989.

[86] C. Hue, J.-P. Le Cadre, and P. Pérez. Tracking multiple objects with particle filtering. *IEEE Transactions on Aerospace and Electronic Systems*, 38(3) :791–812, Jul 2002.

[87] V. S. Hwang. Tracking feature points in time-varying images using an opportunistic selection approach. *Pattern Recognition*, 22(3) :247–256, 1999.

[88] A. Hyvarinen, J. Karhunen, and E. Oja. *Independent Component Analysis.* Wiley-Interscience, 2001.

[89] A. Akhmanova W. J. Niessen E. Meijering I. Smal, I. Grigoriev. Accurate estimation of microtubule dynamics using kymographs and variable-rate particle filters. In X. Pan G. A. Worrell B. He, Y. Kim, editor, *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2009) Held in Minneapolis*, pages 1012–1015, September 2009.

[90] Volocity Improvision.

[91] M. Isard and A. Blake. Condensation – conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1) :5–28, 1998.

[92] P. Jaccard. étude comparative de la distribution florale dans une portion des alpes et des jura. *Bulletin de la Société Vaudoise des Sciences Naturelles*, (37) :547–579, 1901.

[93] A. K. Jain. *Fundamentals of digital image processing.* Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989.

[94] K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. L. L. Schmid, and G. Danuser. Robust single-particle tracking in live-cell time-lapse sequences. *Nature methods*, 5(8) :695 – 702, July 2008.

[95] E. T. Jaynes. *Probability Theory : The Logic of Science (Vol 1).* Cambridge University Press, 2003.

[96] S. Jiang, X. Zhou, T. Kirchausen, and S.T. Wong. Detection of molecular particles in live cells via machine learning. *Cytometry Part A*, 71A(8) :563–575, August 2007.

[97] A. M. Johansen and A. Doucet. *A Tutorial on Particle Filtering and Smoothing : Fifteen years Later.* 2008.

[98] I. M. Johnstone. Wavelet shrinkage for correlated data and inverse problems : adaptivity results. Departament of Statics Stanford University, 1997.

[99] R. Jonker and A. Volgenant. A Shortest Augmenting Path Algorithm for Dense and Sparse Linear Assignment Problems. *Computing*, 38 :325–340, 1987.

[100] S. J. Julier and J. K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3) :401–422, Mar 2004.

[101] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME-Journal of Basic Engineering*, 82(D) :35–45, 1960.

[102] M. Kass, A. Witkin, and D. Terzopoulos. Snakes : active contour models. *International Journal of Computer Vision*, 1 :321–331, 1988.

[103] K. Katoh, K. Hammar, P. J. S. Smith, and R. Oldenbourg. Arrangement of radial actin bundles in the growth cone of Aplysia bag cell neurons shows the immediate past history of filopodial behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 96(14) :7928–7931, 1999.

[104] Z. Khan, T. Balch, and F. Dellaert. MCMC-based particle filtering for tracking a variable number of interacting targets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27 :1805–1819, 2005.

[105] M. Klaas, N. de Freitas, and A. Doucet. Toward practical n2 monte carlo : the marginal particle filter. In *UAI*, pages 308–315, 2005.

[106] A. Kong, S. Liu, and W. Wong. Sequential imputations and bayesian missing data problems. *Journal of the American Statistical Association*, 89, 1994.

[107] T. Kurien. Issues in the design of practical multitarget tracking algorithms. In Y. Bar-Shalom, editor, *Multitarget-Multisensor Tracking : Advanced Applications.* Artech House, Norwood, MA, 1990.

[108] M. Lakadamyali, M. J. Rust, H. P. Babcock, and X. Zhuang. Visualizing infection of individual influenza viruses. *Proceedings of the National Academy of Sciences of the United States of America*, 100(16) :9280–9285, 2003.

[109] M. J. Lehmann, N. M. Sherer, C. B. Marks, M. Pypaert, and W. Mothes. Actin- and myosin-driven movement of viruses along filopodia precedes their entry into cells. *J. Cell Biol.*, 170(2) :317–325, 2005.

[110] N. Li and X. R. Li. Target perceivability and its applications. *IEEE Transactions on Signal Processing*, 49(11) :2588–2604, Nov 2001.

[111] N. Li and X. R. Li. Tracker design based on target perceivability. *Aerospace and Electronic Systems, IEEE Transactions on*, 37(1) :214–225, Jan 2001.

[112] X.-R. Li and V. P. Jilkov. Survey of maneuvering target tracking : dynamic models. In O. E. Drummond, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 4048 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 212–235, July 2000.

[113] T. Lindeberg. Feature detection with automatic scale selection. *International Journal on Computer Vision*, 30(2) :79–116, November 1998.

[114] J. Lippincott-Schwartz and G. H. Patterson. Development and Use of Fluorescent Protein Markers in Living Cells. *Science*, 300(5616) :87–91, 2003.

[115] Y.M. Lu and M.N. Do. Multidimensional directional filter banks and surfacelets. *Image Processing, IEEE Transactions on*, 16(4) :918–931, April 2007.

[116] E. Mairey, A. Genovesio, E. Donnadieu, C. Bernard, F. Jaubert, E. Pinard, J. Seylaz, J.-C. Olivo-Marin, X. Nassif, and G. Duménil. Cerebral microcirculation shear stress levels determine neisseria meningitidis attachment sites along the blood-brain barrier. *J Exp Med*, 203(8) :1939–1950, August 2006.

[117] S. Mallat. *A wavelet tour of signal processing.* Academic Press, 1998.

[118] G. J. Mclachlan. *Discriminant Analysis and Statistical Pattern Recognition (Wiley Series in Probability and Statistics).* Wiley-Interscience, August 2004.

[119] E. Meijering, I. Smal, and G. Danuser. Tracking in molecular bioimaging. *Signal Processing Magazine, IEEE*, 23(3) :46–53, May 2006.

[120] M. M. Mhlanga, D. P. Bratu, A. Genovesio, A. Rybarska, N. Chenouard, U. Nehrbass, and J.-C. Olivo-Marin. In vivo colocalisation of oskar mrna and trans-acting proteins revealed by quantitative imaging of the drosophila oocyte. *PLoS ONE*, 4(7) :e6241, 07 2009.

[121] X. Michalet, F. F. Pinaud, L. A. Bentolila, J. M. Tsay, S. Doose, J. J. Li, G. Sundaresan, A. M. Wu, S. S. Gambhir, and S. Weiss. Quantum Dots for Live Cells, in Vivo Imaging, and Diagnostics. *Science*, 307(5709) :538–544, 2005.

[122] K. G. Murty. An algorithm for ranking all the assignments in order of increasing cost. *Operations Research*, 16 :682–687, 1968.

[123] D. Musicki. Track score and target existence. In *Information Fusion, 2006 9th International Conference on*, pages 1–7, July 2006.

[124] D. Musicki and R.J. Evans. Target existence based mht. *IEEE Conference on Decision and Control*, pages 1228–1233, Dec 2005.

[125] S. Nguyen Ngoc, F. Briquet-Laugier, C. Boulin, and J.-C. Olivo-Marin. Adaptive detection for moving biological objects in video microscopy sequences. In *Proceedings of the International Conference on Image Processing*, volume 3, pages 484–487, 1997.

[126] E. Nicolas, N. Chenouard, J.-C. Olivo-Marin, and A. Guichet. A Dual Role for Actin and Microtubule Cytoskeleton in the Transport of Golgi Units from the Nurse Cells to the Oocyte Across Ring Canals. *Molecular Biology of the Cell*, 20(1) :556–568, 2009.

[127] R. J. Ober, S. Ram, and E. S. Ward. Localization accuracy in single-molecule microscopy. *Biophysical Journal*, 86(2) :1185–1200, February 2004.

[128] K. Okuma, A. Taleghani, O. De Freitas, J. J. Little, and D.G. Lowe. A boosted particle filter : Multitarget detection and tracking. In *In ECCV*, pages 28–39, 2004.

[129] J.-C. Olivo-Marin. Extraction of spots in biological images using multiscale products. *Pattern Recognition*, 35(9) :1989–1996, 2002.

[130] G. Peyré, J. Fadili, and J.-L. Starck. In D. Van De Ville, V. K. Goyal, and M. Papadakis, editors, *Proceedings of the SPIE, Wavelets XII*, volume 6701, 2007.

[131] G. Peyré, J. Fadili, and J.-L. Starck. Learning the morphological diversity. 2009. in press.

[132] D.D.-Y. Po and M.N. Do. Directional multiscale modeling of images using the contourlet transform. *Image Processing, IEEE Transactions on*, 15(6) :1610–1620, June 2006.

[133] A. B. Poore and A. J. Robertson. A new lagrangian relaxation based algorithm for a class of multidimensional assignment problems. *Computational Optimization and Applications*, 8(2) :129–150, 1997.

[134] H. Qian. Single-particle tracking : Brownian dynamics of viscoelastic materials. *Biophysical Journal*, 79(1) :137 – 143, 2000.

[135] H. Qian, M. P. Sheetz, and E. L. Elson. Single particle tracking. analysis of diffusion and flow in two-dimensional systems. *Biophysical journal*, 60(4) :910–921, October 1991.

[136] V. Racine, M. Sachse, J. Salermo, V. Fraisier, A. Trubil, and J.-B. Sibarita. Visualization and quantification of vesicle trafficking on a three-dimensional cytoskeleton network in living cells. *Journal of Microscopy*, 225(3) :214–228, march 2006.

[137] K. Rangarajan and M. Shah. Establishing motion correspondence. *CVGIP : Image Underst.*, 54(1) :56–73, 1991.

[138] D. B. Reid. An algorithm for tracking multiple targets. *IEEE Transactions on Automatic Control*, AC-24 :843–854, 1979.

[139] H. Risken. *The Fokker-Planck Equation : Methods of Solutions and Applications.* Springer Series in Synergetics. Springer, 2 edition, September 1996.

[140] B. Ristic, S. Arulampalam, and N. Gordon. *Beyond the Kalman Filter.* Artech House, 2004.

[141] C. P. Robert and G. Casella. *Monte Carlo Statistical Methods.* Springer-Verlag, 1 edition, August 1999.

[142] M. J. Rust, M. Bates, and X. Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nature Methods*, 3(10) :793–796, October 2006.

[143] D. Sage, F.R. Neumann, F. Hediger, S.M. Gasser, and M. Unser. Automatic tracking of individual fluorescence particles : Application to the study of chromosome dynamics. *IEEE Transactions on Image Processing*, 14(9) :1372– 1383, September 2005.

[144] M. J. Saxton. Single-particle tracking : models of directed transport. *Biophysical Journal*, 67(5) :2110 – 2119, november 1994.

[145] M. J. Saxton. Single-particle tracking : connecting the dots. *Nature Methods*, 5(8) :671–672, 2008.

[146] M. J. Saxton and K Jacobson. Single-Particle Tracking : Applications to Membrane Dynamics. *Annu. Rev. Biophys. Biomol. Struct.*, 26 :373–399, 1997.

[147] I. F. Sbalzarini and P. Koumoutsakos. Feature point tracking and trajectory analysis for video imaging in cell biology. *Journal of Structural Biology*, 151 :182–195, 2005.

[148] A. Sergé, N. Bertaux, H. Rigneault, and D. Marguet. Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes. *Nature methods*, 5(8) :687 – 694, July 2008.

[149] J. Serra. *Image Analysis and Mathematical Morphology.* Academic Press, 1982.

[150] J. K. Sethi and R. Jain. Finding trajectories of feature points in a monocular image sequence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9 :56–73, 1987.

[151] J. A. Sethian. *Level set methods and fast marching methods.* Cambridge University Press, 2nd edition edition, 1999.

[152] E. D. Sheets, R. Simson, and K. Jacobson. New insights into membrane dynamics from the analysis of cell surface interactions by physical methods. *Current Opinion in Cell Biology*, 7(5) :707 – 714, 1995.

[153] M. J. Shensa. The discrete wavelet transform : wedding the a trous and mallat algorithms. *IEEE Transactions on Signal Processing*, 40(10) :2464–2482, 1992.

[154] I. Smal. *Particle Filtering Methods for Subcellular Motion Analysis.* PhD thesis, Erasmus MC, 2009.

[155] I. Smal, K. Draegestein, N. Galjart, W. Niessen, and E. Meijering. Particle filtering for multiple object tracking in dynamic fluorescence microscopy images : Application to microtubule growth analysis. *Medical Imaging, IEEE Transactions on*, 27(6) :789–804, June 2008.

[156] I. Smal, M. Loog, W. J. Niessen, and E. Meijering. Quantitative comparison of spot detection methods in live-cell fluorescence microscopy imaging. In *IEEE International Symposium on Biomedical Imaging : From Nano to Macro - ISBI 2009*, pages 1178–1181, 2009.

[157] I. Smal, E. Meijering, K. Draegestein, N. Galjart, I. Grigoriev, A. Akhmanova, M. E. van Royen, A. B. Houtsmuller, and W. J. Niessen. Multiple object tracking in molecular bioimaging by rao-blackwellized marginal particle filtering. *Medical Image Analysis*, 12(6) :764–777, 2008.

[158] I. Smal, W. Niessen, and E. Meijering. Bayesian tracking for fluorescence microscopic imaging. *Biomedical Imaging : Nano to Macro, 2006. 3rd IEEE International Symposium on*, pages 550–553, April 2006.

[159] I. Smal, W. Niessen, and E. Meijering. Particle filtering for multiple object tracking in molecular cell biology. *Nonlinear Statistical Signal Processing Workshop, 2006 IEEE*, pages 129–132, Sept. 2006.

[160] I. Smal, W. J. Niessen, and E. Meijering. A new detection scheme for multiple object tracking in fluorescence microscopy by joint probabilistic data association filtering. In *IEEE International Symposium on Biomedical Imaging : From Nano to Macro - ISBI 2008*, pages 264–267, 2008.

[161] P. Soille. *Morphological Image Analysis.* Springer-Verlag, New York, 1999.

[162] H. W. Sorenson. On the development of practical nonlinear filters. *Information Sciences*, 7 :253–270, 1974.

[163] J.-L. Starck, E. J. Candes, and D. L. Donoho. The curvelet transform for image denoising. *Image Processing, IEEE Transactions on*, 11(6) :670–684, Jun 2002.

[164] J.-L. Starck, M. Elad, and D.L. Donoho. Redundant multiscale transforms and their application for morphological component analysis. *Advances in Imaging and Electron Physics*, 132, 2004.

[165] J.-L. Starck, M. Elad, and D.L. Donoho. Image decomposition via the combination of sparse representations and a variational approach. *Image Processing, IEEE Transactions on*, 14(10) :1570–1582, Oct. 2005.

[166] J.-L. Starck, J. Fadili, and F. Murtagh. The undecimated wavelet decomposition and its reconstruction. *Image Processing, IEEE Transactions on*, 16(2) :297–309, Feb. 2007.

[167] J.-L. Starck, Y. Moudden, J. Bobin, M. Elad, and D.L. Donoho. Morphological component analysis. In *Proceedings of the SPIE conference wavelets*, volume 5914, July 2005.

[168] J.-L. Starck and F. Murtagh. Multiscale Entropy Filtering. Submitted to Signal Processing, 1998.

[169] J.-L. Starck, F. Murtagh, and A. Bijaoui. Multiresolution support applied to image filtering and restoration. *Graph. Models Image Process.*, 57(5) :420–431, 1995.

[170] J.-L. Starck, F. Murtagh, and A. Bijaoui. *Image processing and data analysis : the multiscale approach.* Cambridge University Press, 2000.

[171] Stearns. Stearns. Green fluorescent protein : The green revolution. *current biology*, 5(3) :262–264, March 1995.

[172] C. Steger. An unbiased detector of curvilinear structures. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(2) :113–125, Feb 1998.

[173] D. Thomann, D.R. Rines, P.K. Sorger, and G. Danuser. Automatic fluorescent tag detection in 3D with super-resolution : application to the analysis of chromosome movement. *Journal of Microscopy*, 208(1) :49–64, October 2002.

[174] MetaMorph Imaging System Universal Imaging Corporation.

[175] H. L. Van Trees. *Detection, Estimation and Modulation Theory, Part I.* Wiley, New York, 1968.

[176] J. Vermaak, A. Doucet, and P. Pérez. Maintaining multi-modality through mixture tracking. In *Int. Conf. on Computer Vision, ICCV'03*, Nice, France, June 2003.

[177] J. Vermaak, S. Godsill, and P. Pérez. Monte Carlo filtering for multi-target tracking and data association. *IEEE Trans. on Aerospace and Electronic Systems*, 41(1) :309–332, 2005.

[178] L. Vincent. Morphological gray scale reconstruction in image analysis : Applications and efficient algorithms. *Image Processing, IEEE Transactions on*, 2(2) :176–201, April 1993.

[179] C. Vonesch, F. Aguet, J.-L. Vonesch, and M. Unser. The colored revolution of bioimaging. *IEEE Signal Processing Magazine*, 23(3) :20–31, 2006.

[180] C. M. Waterman-Storer, A. Desai, C. Bulinski, and E. D. Salmon. Fluorescent speckle microscopy, a method to visualize the dynamics of protein assemblies in living cells. *Current Biology*, 8(22) :1227–S1, November 1998.

[181] G. A. Watson and W. D. Blair. Imm algorithms for tracking targets that maneuver through coordinated turns. *SPIE*, 1698 :236–247, 1992.

[182] G. H. Weiss and R. J. Rubin. Random walks : theory and selected applications. *Adv. Chem. Phys.*, 52 :363–505, 1983.

[183] P. Willet, Y. Ruan, and R. Streit. PMHT : problems and some solutions. *IEEE Transactions on Aerospace and Electronic Systems*, 38(3) :738–754, 2002.

[184] A. Woiselle, J.-L. Starck, and J. Fadili. 3D inpainting using sparse representations. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7446 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, August 2009.

[185] L. Ying, L. Demanet, and E. Candes. In M. Papadakis, A. F. Laine, and M. Unser, editors, *Proceedings of the SPIE, Wavelets XI*, volume 5914, septembre 2005.

[186] B. Zhang, J. Fadili, J.-L. Starck, and J.-C. Olivo-Marin. Multiscale variance-stabilizing transform for mixed-poisson-gaussian processes and its applications in bioimaging. In *Proc. IEEE International Conference on Image Processing ICIP 2007*, volume 6, pages VI–233–VI–236, Sept. 16 2007–Oct. 19 2007.

[187] B. Zhang, M.J. Fadili, J.-L. Starck, and S.W. Digel. Fast poisson noise removal by biorthogonal haar domain hypothesis testing. *Statistical Methodology*, 5(4) :387 – 396, 2008. Astrostatistics.

[188] B. Zhang, J. Zerubia, and J.-C. Olivo-Marin. Gaussian approximations of fluorescence microscope point-spread function models. *Applied Optics*, 46(10) :1819–1829, April 2007.

[189] C. Zimmer, B. Zhang, A.C. Dufour, A. Thebaud, S. Berlemont, V. Meas-Yedid, and J.-C. Olivo-Marin. On the digital trail of mobile cells. *Signal Processing Magazine*, 23(3) :54–62, May 2006.