



HAL
open science

Modèles statistiques pour l'analyse d'images, algorithmes et applications en imagerie bio-cellulaire et moléculaire

Charles Kervrann

► **To cite this version:**

Charles Kervrann. Modèles statistiques pour l'analyse d'images, algorithmes et applications en imagerie bio-cellulaire et moléculaire. Interface homme-machine [cs.HC]. Université Rennes 1, 2010. tel-00541298

HAL Id: tel-00541298

<https://theses.hal.science/tel-00541298>

Submitted on 30 Nov 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HABILITATION À DIRIGER DES RECHERCHES

présentée devant

L'Université de Rennes 1
Spécialité : Traitement du Signal et Télécommunications

par

Charles Kervrann

**Modèles statistiques pour l'analyse d'images, algorithmes et applications
en imagerie bio-cellulaire et moléculaire**

soutenue le 29 juin 2010 devant le jury composé de :

M.	Gilles Celeux	Président
MM	Nicholas Ayache	Rapporteur
	Jean-Michel Morel	Rapporteur
	Christoph Schnörr	Rapporteur
MM	Rémi Gribonval	Examineur
	Jean Salamero	Examineur

à Mary, Marianne, Hugo, Justine.
à ma famille.

à la mémoire de Jean-Pierre LC.

Table des matières

Remerciements	11
Avant-propos	13
1 Introduction	17
2 Modèles statistiques et variationnels pour la segmentation d'images	23
2.1 Modèles bayésiens	23
2.1.1 Champs de Markov et champs aléatoires de Gibbs	23
2.1.2 Approche bayésienne et sélection de modèles	25
2.1.3 Segmentation bayésienne d'images	26
2.2 Modèles variationnels	31
2.2.1 Minimisation de la variation totale	32
2.2.2 Segmentation d'images et équations aux dérivées partielles	34
2.3 Un principe morphologique, statistique et variationnel pour la segmentation . . .	35
2.3.1 Représentation de l'image par ses lignes de niveau	36
2.3.2 Segmentation par minimisation de fonctionnelles d'énergie et sélection de lignes de niveau	37
2.3.3 Un autre éclairage : sélection de modèles, représentation parcimonieuse, modèles a contrario	43
2.4 Conclusion et perspectives	47
3 Approches non-paramétriques pour l'analyse d'images	49
3.1 Myopie, redondance et modèles d'images	49
3.2 Estimation non-paramétrique pour le débruitage d'images	50
3.2.1 Filtres de voisinage et méthodes non-linéaires	51
3.2.2 Estimation adaptative et équilibre biais-variance	52

Table des matières

3.2.3	Un autre éclairage statistique	58
3.2.4	Deux exemples d'application des méthodes adaptatives	62
3.3	Détection non-paramétrique de changements dans une paire d'images	67
3.3.1	Le détecteur classique en traitement du signal	68
3.3.2	Principe de la fusion de décisions locales	70
3.3.3	Approche non-paramétrique et multi-échelle	71
3.4	Conclusion et perspectives	75
4	Modèles et algorithmes en action : analyse du transport intra-cellulaire	79
4.1	La vidéo-microscopie : de nouveaux défis en analyse d'images	80
4.2	Débruitage spatio-temporel en vidéo-microscopie	82
4.3	Correction du photo-blanchiment et soustraction de « fond »	87
4.3.1	Approche spatio-temporelle et semi-paramétrique	87
4.3.2	Approche temporelle et non-paramétrique	91
4.3.3	Bilan	94
4.4	Détection de changements spatio-temporels en vidéo-microscopie	95
4.4.1	Sélection de modèles unidimensionnels	95
4.4.2	Analyse semi-locale des occurrences dans une paire d'images	97
4.5	Estimation de trafic par « Tomographie de Réseaux »	101
4.6	Conclusion	109
5	Quelques perspectives	111

Glossaire (microscopie)

Cryo-ET (Cryo-Electron Tomography) : représentation 3D à partir de projections 2D d'objets sub-cellulaires et moléculaires (figés à très basse température) à la résolution de 5-20 nm à l'aide d'une microscope électronique à transmission (un faisceau d'électrons traverse l'échantillon et produit une vue 2D en projection).

FLIM (Fluorescence Lifetime Microscopy Imaging) : imagerie microscopique de durée de vie de fluorescence (quelques nano-secondes).

FRAP (Fluorescence Recovery After Photobleaching) : analyse de la dynamique cellulaire par des techniques de retour de fluorescence après photo-extinction locale.

FRET (Förster Resonance Energy Transfer) : microscopie pour mesurer le transfert d'énergie entre molécules fluorescentes (interactions protéiques).

MSDC (Multipoint Spinning Disk Confocal) : microscopie confocale (illumination focalisée (laser)) rapide multipoints et disques rotatifs.

PALM (Photo-Activated Localization Microscopy) : microscopie de superrésolution par photo-activation stochastique de fluorophores et ajustement de fonctions de transfert optique (Betzig, Patterson, Sougrat, Lindwasser, Olenych, Bonifacino, Davidson, Lippincott-Schwartz & Hess 2006).

SIM (Structured Illumination Microscopy) : microscopie superrésolution (type « champ-large ») avec illumination contrainte par des motifs structurés et analyse des interférences (formes de Moiré) (Gustafsson, Shao, Carlton, Wang, Golubovskaya, Agard & Sedat 2008).

TIRF (Total Internal Reflectance Microscopy) : microscopie 2D avec illumination à ondes évanescentes (Axelrod 1981).

WFM (Wide-Field Microscopy) : microscopie optique avec illumination non focalisée de l'échantillon.

Table des matières

Remerciements

JE tiens d'abord à remercier les membres du jury, en particulier Christoph Schnörr, Professeur à l'université de Heidelberg, Nicholas Ayache, Directeur de Recherches à l'Inria Sophia-Antipolis et Jean-Michel Morel, Professeur à l'École Normale Supérieure de Cachan, qui ont accepté la lourde tâche d'évaluer ces travaux.

Je remercie ensuite Gilles Celeux, Directeur de Recherches à l'Inria Saclay qui a accepté de présider ce jury, Rémi Gribonval, Directeur de Recherches à l'Inria Rennes Bretagne-Atlantique, et Jean Salamero, Directeur de Recherches Cnrs à l'Institut Curie pour leur participation à ce jury.

J'adresse mes plus vifs remerciements à Alain Trubuil, avec qui j'ai eu la chance de collaborer ces dix dernières années, pour ses conseils amicaux et avisés, une exemplarité en terme de qualités humaines, de méthodes de travail et de démarches scientifiques. Certaines idées sont à mettre à son crédit et ont servi de terreau aux recherches que j'ai menées par la suite. Merci à Jérôme Boulanger, Thierry Pécot et Sophie Blestel, qui reconnaîtront dans ce manuscrit des aspects de leur travail de thèse, pour m'avoir écouté ces dernières années et avec lesquels j'ai eu tant de plaisir à discuter.

Merci encore et surtout à tous les membres de l'unité Inra MIA (Mathématiques et Informatique Appliquées) de Jouy-en-Josas et à ceux de l'équipe-projet Vista (VIsion-Spatio Temporelle et Apprentissage) de l'Inria Rennes, stagiaires, doctorants, ingénieurs, assistantes, chercheurs, avec qui j'ai eu le plaisir de collaborer toutes ces années à Jouy-en-Josas puis à Rennes ; j'ai une pensée particulière pour Elisabeth de Turckheim qui, avec enthousiasme, m'a fait confiance lors de ma demande de « mise à disposition » auprès de l'Inria, et Bruno Goffinet et Patrick Bouthemy qui, avec bienveillance, m'ont soutenu et ont réussi à maintenir une relation de confiance entre l'Inra et l'Inria.

Je tiens enfin à remercier ici Jean Salamero, qui représente bien plus qu'un collaborateur désormais, et les nombreux collègues et doctorants de l'Inria Rennes & Irisa (équipes Fluminance, Lagadic, Temics, Visages), de l'Umr 144 Cnrs Institut Curie (Plateforme Imagerie Tissulaire et Cellulaire IBiSA), de l'Umr 6026 Cnrs (équipe TIPS) et le bureau GdR 2588 Cnrs (Microscopie

Remerciements

Fonctionnelle du Vivant) avec qui j'ai eu le plaisir de travailler :

Sabine Bardin	Fabrice Heitz
Christian Barillot	Pierre Hellier
Ikhlef Béchar	Mark Hoebeke
Huguette Béchu	Hervé Jégou
Sophie Blestel	Bertrand Kaeffer
Jérôme Boulanger	Kiên Kiêu
Frédéric Cao	Claude Labit
Bruno Cernuschi-Frias	Ivan Laptev
Anatole Chessel	Jean-Pierre Le Cadre
Denis Chrétien	David Legland
Bertrand Cinquin	Etienne Mémin
Thomas Corpetti	Arnaud Ogier
Pierrick Coupé	Lissia Pardini
Bernard Delyon	Thierry Pécot
Frank Davoine	Patrick Pérez
Marie-Francoise Devaux	Mathieu Pinot
Ronan Fablet	Sylvain Prima
Zoher Guerroui	Victor Racine
Alexandre Gidon	John Sedat
Charles Gueudry	Jean-Baptiste Sibarita
Bruno Goud	Francois Waharte
Patrick Héas	Jian-Feng Yao

J'espère n'avoir oublié personne et je tiens, par avance, à m'excuser auprès de ceux qui auraient mérités d'être cités aussi. Je remercie Sylvain et Sophie qui ont consacré quelques heures à me lire attentivement ...

Les travaux de recherche décrits dans ce mémoire ont été menés entre 1998 et 2003, d'abord dans l'unité Inra MIA à Jouy-en-Josas puis, à compter de 2003, dans l'équipe-projet Vista à l'Inria Rennes & IriSa. Sur cette période, j'ai essentiellement porté mes efforts sur l'analyse des images 2D-3D et sur le traitement des séquences temporelles d'images, à l'aide de méthodes statistiques/probabilistes ou variationnelles (minimisation de fonctionnelles) que j'exposerai dans ce mémoire.

Mon activité scientifique relève du traitement d'images, de l'algorithmique et des mathématiques appliquées. Même si les approches développées se veulent génériques, les applications sont très souvent présentes dans mon quotidien. Les choix méthodologiques que j'ai pu faire ont été généralement guidés par le domaine d'application. Ces dernières années, j'ai privilégié le développement de méthodes d'analyse d'images dans un contexte scientifique important, celui de la biologie cellulaire et de l'imagerie de fluorescence. La modélisation statistique, thème que que j'ai surtout traité, a permis d'aborder très efficacement certains problèmes de traitement d'images. Reconnue pour son efficacité en traitement d'images, elle soulève toujours des questions scientifiques passionnantes.

Dans le contexte pluridisciplinaire où je me suis placé, le dialogue entre les communautés des mathématiques appliquées, des statistiques, du traitement du signal et de l'image, puis de la biologie cellulaire et de la physique (bio-physique, optique, ...), ouvre des perspectives scientifiques nouvelles, parfois inattendues, et culturellement enrichissantes. J'espère pouvoir démontrer dans ce mémoire, via quelques contributions méthodologiques, les avantages de cette démarche qui réclame une ouverture d'esprit et une certaine curiosité sur des savoir-faire développés dans des communautés scientifiques voisines ou dans d'autres disciplines scientifiques telles que la biologie cellulaire. Les avancées de l'état des connaissances dans les « Sciences du Vivant » d'une part, et pour ma propre discipline d'autre part, nécessitent un dialogue étroit, et donc un investissement personnel pour comprendre ses interlocuteurs qui ne maîtrisent pas toujours le même langage.

J'ai rejoint l'Inra en qualité de chargé de recherche 2ème classe le 1er novembre 1997, pour intégrer l'unité MIA (Mathématique et Informatique Appliquées) sur le centre de recherches de Jouy-en-Josas. Au cours de ces années, j'ai engagé un programme de recherche en « imagerie biologique » qui paraissait quelque peu éloigné des objectifs finaux de l'Institut et des thématiques principales du département MIA. Tout en développant ces activités en analyse d'images, je me suis familiarisé avec une partie des statistiques (surtout non-bayésiennes) que j'ignorais dans une ambiance chaleureuse et toujours très amicale. J'ai surtout eu la chance de rencontrer Alain Trubuil avec qui j'ai effectué un long chemin, partageant des réussites et quelques déceptions. À

compter du 1er juillet 2003, j'ai souhaité être « mis à disposition » auprès de l'Inria à l'Irisa (Institut de la Recherche en Informatique et Systèmes Aléatoires) dans l'équipe-projet Vista (VIsion Spatio-Temporelle et Apprentissage) dont la spécificité est le traitement de séquences temporelles d'images pour des applications variées : vidéo, imagerie météorologique, imagerie médicale. Les thèmes méthodologiques de l'équipe-projet Vista sont étroitement liés à l'estimation statistique, aux techniques de filtrage temporel stochastique (filtrage « particulière »), aux modèles spatiaux et markoviens, aux modèles probabilistes pour la reconnaissance de formes et aux modèles de mécanique des fluides. Dans cette équipe, dirigée par Patrick Bouthemy, j'ai poursuivi mes activités et développé des méthodes originales et adaptées au traitement de séquences d'images en vidéo-microscopie pour la biologie cellulaire et moléculaire. Grâce notamment aux projets MoDynCell5D (ACI IMPBio 2004-2007) et DynaMIT (ARC Inria 2007-2008 et AIP Inra-Inria 2009-2010), j'ai mesuré l'importance de collaborer étroitement sur cette thématique avec des équipes de recherche de pointe en biologie cellulaire telle que celles de l'Umr 144 Cnrs (Jean Salamero et Jean-Baptiste Sibarita) de l'Institut Curie et de l'Umr 6026 de l'université de Rennes 1 (Denis Chrétien). En encadrant à la même époque Jérôme Boulanger, Thierry Pécot, Anatole Chessel (post-doc) et Sophie Blestel, j'ai essayé d'aller plus loin dans la résolution des problèmes que nous rencontrions. J'espère avoir réussi à leur communiquer, dans le cadre de notre collaboration, une certaine passion du métier de chercheur. Celle-ci m'avait été transmise par Fabrice Heitz à l'Irisa/Inria Rennes, quelques années plus tôt dans l'équipe-projet Temis dirigée alors par Claude Labit.

L'analyse de problématiques dans un domaine applicatif permet bien souvent d'exhiber des problèmes originaux pour lesquels les solutions restent à inventer. À l'inverse, la mise en perspective de ces problématiques sous une forme mathématique ou algorithmique permet souvent de s'affranchir de l'application visée. *In fine*, de nouvelles méthodologies génériques peuvent voir le jour, à la fois efficaces pour l'application visée et de manière surprenante, pour d'autres problèmes « similaires ». Ce mémoire est donc organisé en quatre chapitres principaux. Après une brève introduction, je décris dans le chapitre 2 quelques modélisations mathématiques pour l'analyse d'images et quelques solutions aux « problèmes inverses » posés. Ce chapitre constitue une part importante et la plus personnelle de mon activité scientifique au sein de l'unité Inra MIA de Jouy-en-Josas. Le chapitre suivant est consacré aux problèmes de la restauration (ou débruitage) d'images et à la détection de changements dans des séries temporelles d'images. Il s'agit de deux volets de mon activité auxquels ont contribué significativement plusieurs doctorants et collaborateurs. Je décris avec quelques détails ces approches qui exploitent largement les primitives « motifs » (ou « *patches* »), avec une efficacité plus importante que celle que j'avais considérée au tout début de ma thèse de doctorat (segmentation d'images texturées (Kervrann & Heitz 1994, Kervrann & Heitz 1995)). Les algorithmes précédents procèdent bien souvent de manière très locale, et il peut sembler plus naturel d'améliorer la « myopie » de ces algorithmes en considérant des primitives du type « motifs » dans des voisinages spatiaux de taille variable. Les méthodes en question, assez génériques, sont établies dans le cadre applicatif de l'imagerie bio-cellulaire (voir chapitre 4) mais se sont avérées également très pertinentes pour d'autres domaines applicatifs : vidéo (Boulanger, Kervrann & Bouthemy 2007), imagerie médicale (Coupé, Hellier, Kervrann & Barillot 2009, Coupé, Hellier, Prima, Kervrann & Barillot 2008, Coupé, Yger, Prima, Hellier, Kervrann & Barillot 2008)... Le chapitre 4 est aussi consacré à la modélisation du trafic vésiculaire et du transport membranaire à l'aide de techniques de « Tomographie

de Réseaux » (Vardi 1996), peu répandues dans les communautés du traitement d'images et de la vision par ordinateur. Les méthodes principalement non-paramétriques présentées dans les chapitres 2 et 3 trouvent aussi leur intérêt dans cette tentative de modélisation « sans trop d'*a priori* » des images acquises en microscopie et biologie cellulaire. Enfin, je conclurai dans un dernier chapitre par quelques réflexions personnelles et je présenterai les axes de recherche que je souhaite développer dans les prochaines années.

CHAPITRE 1

Introduction

Analyser et traiter les images numériques afin d'extraire des informations quantitatives a suscité l'intérêt d'une très large communauté de chercheurs depuis longtemps. L'augmentation très nette du nombre de journaux et conférences ces dernières années témoigne de cette détermination et de cet engouement. Néanmoins, ce constat met aussi en évidence les difficultés rencontrées. Plusieurs échecs, mais aussi quelques succès, ont émaillé cette période où l'informatique, le traitement du signal et la théorie de l'information, les mathématiques appliquées, les statistiques, la théorie des probabilités, la géométrie et la reconnaissance des formes... ont joué un rôle moteur à un moment ou à un autre durant cette période qui a débuté dans les années 70.

Au cours de la période la plus récente, on peut remarquer que l'optimisation discrète et continue semblent être arrivées à maturité. Les méthodes de « *graph-cuts* » (ou « *min-cut/max-flow* ») (Greig, Porteous & Seheult 1989, Boykov & Jolly 2001, Boykov, Veksler & Zabih 2001, Kolmogorov & Zabih 2001, Boykov & Kolmogorov 2003) ou d'ensembles de niveau (« *level sets* ») (Osher & Sethian 1988, Caselles, Kimmel & Sapiro 1995, Sethian 1996, Caselles, Kimmel & Sapiro 1997, Schnörr 1998, Paragios & Deriche 2000) établies dans le cadre des équations aux dérivées partielles (EDPs), garantissent le minimum global d'une fonction objective qui correspond à la modélisation d'un problème de segmentation ou de recalage d'images. La représentation d'images par méthodes de décomposition en ondelettes (Donoho & Johnstone 1994, Mallat 1999, Starck, Candes & Donoho 2000, Le Pennec & Mallat 2005) ou méthodes parcimonieuses (Elad & Aharon 2006c, Mairal, Sapiro & Elad 2008) d'une part, et d'autre part, la restauration d'images par minimisation de la variation totale (Rudin, Osher & Fatemi 1992) et équations aux dérivées partielles (Perona & Malik 1990a, Alvarez, Lions & Morel 1992, Caselles, Morel, Sapiro & Tannenbaum 1998, Weickert 1999a, Tschumperlé & Deriche 2005, Aujol, Aubert, Blanc-Féraud & Chambolle 2005) figurent parmi les succès notablement reconnus. Dans le domaine des statistiques et de la théorie de l'apprentissage, les méthodes à noyaux (Schölkopf, Tsuda & Vert 2004), les machines à vecteurs de supports (SVM) ou les méthodes de « *boosting* » (Viola & Jones 2001),

font désormais partie de la panoplie des méthodes de classification et de reconnaissance de formes.

Néanmoins, il est assez frappant de remarquer que la modélisation d'images a davantage été développée au milieu des années 80, une période où la modélisation par champs aléatoires et modèles de Markov (Geman & Geman 1984, Blake & Zisserman 1987, Leclerc 1989, Morel & Solimini 1994, Mumford 1994) connut un grand succès. Les méthodes variationnelles de segmentation d'images (Mumford & Shah 1989, Kass, Witkin & Terzopoulos 1987, Cohen 1991, Cohen, Cohen & Ayache 1992, Caselles et al. 1995, Zhu & Yuille 1996) connurent un succès sans précédent également. Paradoxalement, c'est assez tardivement que des algorithmes de minimisation des critères de segmentation furent proposés. Jusqu'à la fin des années 80, les connaissances *a priori* s'exprimaient encore sous la forme de potentiels quadratiques, davantage que pour des questions de mise en œuvre dans des schémas numériques, pour des questions de modélisation réaliste. Le *Traitement du Signal* (filtre de Wiener, filtre de Kalman, ...) était alors l'outil de référence en la matière. Les méthodes non-linéaires étaient principalement proposées par la communauté de morphologie mathématique (Serra 1982, Vincent & Soille 1991, Salembier & Serra 1995), dans un cadre de modélisation discrète.

De manière assez déroutante mais finalement très stimulante, beaucoup de chemins ont été parcourus depuis par les communautés de traitement d'images et de vision par ordinateur. Toutefois, il est désormais de moins en moins rare de voir se constituer des groupes de recherche composés de physiciens pour décrire les modèles de formation d'images, de mathématiciens appliqués, de statisticiens, de probabilistes, d'informaticiens voire de biologistes ou de médecins pour analyser les informations disposées sur un support numérique ou une grille de pixels. Les progrès pour des disciplines majeures et fondamentales telles que la biologie, la chimie et la physique sont désormais des enjeux importants pour la décennie à venir, à la fois pour la communauté du traitement de l'image mais aussi pour celles des mathématiques appliquées et de l'informatique.

Imagerie physique et quantification

En imagerie optique, il est bien établi que l'intensité lumineuse enregistrée sur une grille de pixel (carré de côté 1) est une quantification de la distribution de lumière observée. Cette quantification peut être plus ou moins fine selon le type d'observations envisagé. Une quantification sur 256 niveaux est souvent suffisante pour interpréter une scène de notre monde environnant par la vision humaine. Dans des domaines plus proches de la physique, l'astronomie, la télédétection, l'imagerie médicale ou la microscopie, une quantification sur 16 bits ou 24 bits est bien souvent nécessaire pour une meilleure classification des informations. En imagerie photonique, l'intensité lumineuse enregistrée est proportionnelle au nombre de photons mesurés. Dans des cas limites, il peut ne pas excéder 1 à 2 photons par pixel, rendant impossible l'interprétation par l'oeil humain. Dans ce document, nous nous intéressons à ce type d'imagerie et au cas plus général des images scalaires pour lequel nous proposons des modélisations et des interprétations.

Traitement du signal et/ou analyse des formes

À l'origine, l'image a été perçue comme une extension naturelle à la dimension 2 de signaux aléatoires. Les outils de traitement du signal et l'analyse dans le domaine de Fourier se sont avérés pertinents, notamment pour traiter des signaux stationnaires. Une grande partie des ré-

sultats théoriques obtenus pour des signaux se transposent sans difficulté aux images. Les notions d'échantillonnage, de corrélation, de filtrage linéaire ... sont très bien adaptées. Néanmoins, la notion de causalité propre aux signaux n'est pas naturelle en imagerie. D'emblée, on peut en effet remarquer que la plupart des images 2D que nous manipulons correspondent à la projection de scènes 3D sur un plan. Les objets sont géométriquement remarquables et l'analyse des formes est essentielle pour faciliter la reconnaissance. Il est donc nécessaire de proposer des modèles et des représentations adaptées à la description de ces objets projetés. Une modélisation précise de la projection de la scène 3D et le recalage de ces modèles doit permettre une bonne interprétation de la scène observée.

Une modélisation fidèle a néanmoins des conséquences importantes dans les schémas d'estimation. La théorie des champs de Markov, distincte sur certains aspects de la théorie des chaînes de Markov, est considérée comme une avancée majeure et déterminante sur ce point. Elle permet de prendre en compte la géométrie dans les schémas d'analyse, une notion qui échappait à la théorie du traitement du signal étendue aux images. L'analyse par ondelettes géométriques s'inscrit dans la continuité de cette évolution, menant à des algorithmes de débruitage et de compression de l'information très performants. Pourtant, malgré ces efforts et d'autres tout aussi importants durant plusieurs décennies, il apparaît clairement que les modèles d'images retenus sont bien souvent une représentation idéale et très simplifiée.

Dans les faits, plusieurs alternatives s'offrent au traiteur d'images. Il peut proposer une modélisation discrète et des algorithmes pour extraire l'information perçue. Dans certains contextes, l'image est davantage une réalisation d'un vecteur aléatoire. Cette abstraction est possible dans la mesure où l'image numérique est un objet discret relativement limité (une matrice de nombres). Le traiteur d'images est également tenté par une modélisation plus complexe qui correspond à un phénomène physique plus proche de son intuition et de la réalité traduite par l'image. Dans ce cas, une modélisation continue et tridimensionnelle semblera plus appropriée. Quelle modélisation choisir et quel avantage procure un choix plutôt qu'un autre ? Probablement, les cadres de modélisation les plus pertinents correspondent à des choix pragmatiques. Ils décrivent bien la réalité physique, ou bien permettent la mise en œuvre d'algorithmes permettant d'extraire une information utile tout en s'appuyant sur une modélisation relativement simpliste du phénomène étudié. Finalement, un compromis est bien souvent choisi, satisfaisant son intuition et permettant la mise en œuvre d'algorithmes numériques, parfois avec une nécessité d'obtenir une réponse très rapidement.

Effectivement, la résolution numérique est tout aussi importante que la modélisation elle-même. Dans la période la plus récente, cet aspect a été largement développé au détriment de la modélisation. Les algorithmes numériques qui visent à optimiser des paramètres ou atteindre le minimum d'un critère objectif dans un temps limité sont essentiels pour un grand nombre d'applications. La résolution numérique d'équations aux dérivées partielles résultant d'une modélisation continue ne pourrait pas se faire sans une recherche appropriée sur cette question. A titre d'exemple, manipuler des courbes pour extraire des objets d'intérêt de manière souple en tenant compte des changements de topologie n'est finalement arrivé qu'au milieu des années 90. Des modélisations très complexes (Mumford & Shah 1989) n'ont ainsi pu être réellement tentées qu'au prix de ces efforts issus d'autres disciplines. Le constat est identique pour la modélisation par champs de Markov qui connaît ces dernières années un regain de forme suite aux avancées en optimisation et théorie des graphes. Donc, bien souvent, ces réflexions en termes d'optimi-

sation ou de modélisation (déterministe ou stochastique) n'ont pu être possibles que par une connaissance approfondie de résultats en mathématiques appliquées et probabilités.

Restauration et segmentation d'images

Les modélisations très générales que nous venons d'évoquer jusqu'à présent se définissent d'abord dans une logique de traitement, au sens où il s'agit de calculer une nouvelle image suite à quelques opérations menées sur l'image originale. Extraire une information est une tâche bien plus complexe et a fait l'objet de nombreux travaux ces dernières décennies. La vision artificielle quant à elle réclame des modèles qui calculent des objets du monde réel. Il s'agit d'analyser des images en relation avec le monde réel physique qu'elles représentent afin de faciliter une prise de décision ou une reconnaissance. Une part de mon activité de recherche a donc aussi été marquée par ma contribution au problème fondamental de la segmentation d'images (2D et 3D). On rappelle que la segmentation d'images consiste à produire une partition de l'image initiale en régions homogènes (e.g. au sens de la luminance, texture ou du mouvement), conduisant aux différentes composantes de la scène. L'analyse bayésienne d'images (Geman & Geman 1984, Besag 1986, Geiger & Yuille 1991, Grenander & Miller 1994) et l'analyse variationnelle (Mumford & Shah 1989, Kass et al. 1987, Morel & Solimini 1994, Caselles et al. 1995, Schnörr 1998, Heiler & Schnörr 2005) procurent des solutions élégantes à ces problèmes de segmentation. Elles mettent en jeu des variables discrètes (modèles markoviens) ou continues et se traduisent, en définitive, par la formulation d'un problème de minimisation globale d'une fonction objective ou d'une énergie. L'énergie à minimiser se décompose généralement en deux termes distincts : un premier terme d'attache aux données (contraste) mesurant la relation entre le modèle de l'image et les observations et un second terme de régularisation, pouvant favoriser l'émergence de régions homogènes délimitées par des bords réguliers, d'un faible nombre de régions ou d'objets de forme géométrique ou textuelle voire de mouvement, connus *a priori*. À la différence de la restauration d'images qui cherche une fonction sur une grille discrète ou continue, la segmentation est aussi basée sur des notions de géométrie puisqu'il s'agit d'extraire des objets dont la forme est connue plus ou moins précisément. Parmi l'ensemble des démarches possibles, on peut tirer avantage de la description de l'image par ses lignes de niveau introduit par Matheron (1975). Grâce à ces lignes de niveau (Ballester, Cubero-Castan, Gonzalez & Morel 2000), on peut doter l'image d'un vrai modèle géométrique. En convoluant l'image « faiblement », il est alors possible de s'affranchir quelque peu du problème de la discrétisation du support de l'image. Les primitives à considérer sont désormais ces courbes et non plus des pixels agrégés pour d'écrire des régions.

Un problème moins « mal posé » mais pas davantage résolu à vrai dire est celui de la restauration ou débruitage d'images. En traitement du signal, les méthodes qui s'appuient sur une représentation par ondelettes des signaux 2D, ont fait leurs preuves pour améliorer la qualité des images, c'est-à-dire supprimer le bruit introduit par les capteurs. Cette composante aléatoire peut être considérée comme stationnaire ou non selon les cas. Ce problème de la restauration d'images est également étroitement lié à celui de la représentation puisqu'il s'agit d'annuler les coefficients non significatifs ou d'estimer un modèle paramétrique de complexité variable (e.g. polynôme) dans un voisinage limité. Des résultats très récents en analyse harmonique, représentation parcimonieuse et analyse variationnelle prouvent que ce problème de la représentation est toujours en effervescence. Par ailleurs, assez éloignées, les communautés des ondelettes et

des EDPs tendent désormais à se rapprocher. Le point de convergence : la géométrie exprimée par des ondelettes géométriques (Starck et al. 2000, Le Pennec & Mallat 2005), des tenseurs de structure (Weickert 1999*a*), des représentations « espace-échelle » (Alvarez, Guichard, Lions & Morel 1993). Les critères employés et les motivations sont aussi cohérents avec une modélisation markovienne des images.

Modèles statistiques et variationnels pour la segmentation d'images

Les outils statistique et probabiliste sont sûrement les mieux adaptés pour caractériser la loi des images. Néanmoins, dans le cadre plus restreint de la résolution de « problèmes inverses », les estimateurs bayésiens au sens du maximum *a posteriori* (MAP) ont des propriétés communes avec les modèles variationnels établis dans un cadre déterministe. Nous décrirons ces liens dans ce chapitre consacré à la segmentation d'images.

2.1 Modèles bayésiens

L'inférence statistique est une discipline qui a déjà fait ses preuves par le passé (Besag 1974, Geman & Geman 1984, Besag 1986) en traitement d'images. Disposer d'échantillons de grande taille pour l'apprentissage n'est pas un problème en imagerie compte tenu de la quantité d'images ou de films acquis chaque jour. Le nombre de pixels qui composent une image est aussi très important. Il est donc possible d'extraire des informations sur la nature statistique des éléments qui composent les images. Nous verrons par dans ce mémoire, à l'instar de Lee, Pedersen & Mumford (2003) et Koloydenko & Geman (2006), que le nombre de motifs de petite taille au sein d'une image est suffisamment élevé pour caractériser les « formes » textuelles et géométriques qu'elle contient.

2.1.1 Champs de Markov et champs aléatoires de Gibbs

Soit $\mathbf{u} = (u_1, \dots, u_N)$ la réalisation d'un vecteur aléatoire où $u_s \in \Lambda_s$ et Λ_s désigne l'espace d'états au site s . On note $\mathcal{S} = \{1, \dots, N\}$ l'ensemble fini de sites $s \in \mathcal{S}$ et $\mathcal{U} = \prod_{s=1}^N \Lambda_s$ l'ensemble des configurations. Dans le cas simple où chaque variable u_s prend ses valeurs dans le même espace d'états Λ quel que soit $s \in \mathcal{S}$, alors $\mathbf{u} \in \mathcal{U} = \Lambda^N$.

Dans un cadre discret, les champs aléatoires de Gibbs sont des modèles probabilistes pour lesquels les densités de probabilité s'expriment usuellement de la manière suivante :

$$p(\mathbf{u}) = \frac{1}{Z} \exp(-E(\mathbf{u})) = \frac{1}{Z} \exp\left(-\sum_{c \in \mathcal{C}} E_c(\mathbf{u})\right) \quad (2.1)$$

où $E(\mathbf{u}) = \sum_{c \in \mathcal{C}} E_c(\mathbf{u})$ représente l'énergie, $E_c(\mathbf{u})$ désigne le potentiel d'« interaction locale » et Z est la constante de normalisation assurant que la somme sur l'ensemble de toutes configurations de la densité est égale à 1. Une clique c est un sous-ensemble (petit paquet d'indices) de \mathcal{S} dont les éléments sont mutuellement voisins. Nous notons également \mathcal{C} l'ensemble de toutes les cliques du graphe \mathcal{G} qui permet de décrire les interactions locales sous-jacentes. On considère donc un graphe non-dirigé $\mathcal{G} = (\mathcal{A}, \mathcal{V})$ défini sur \mathcal{S} où \mathcal{V} représente l'ensemble des sommets et \mathcal{A} l'ensemble des arêtes du graphe. Les sites s et t sont dits voisins (soit encore $s \sim t$) si $(s, t) \in \mathcal{A}$, ce qui revient à supposer qu'il existe une arête entre s et t . On considère une relation symétrique de voisinage entre les sites s et t ; par convention, le site s n'est pas voisin de lui-même.

Si la densité de probabilité $p(\cdot)$ d'un champ de Markov est partout positive, alors le théorème d'Hammersley-Clifford établit l'existence d'une représentation par champs aléatoires de Gibbs de ce champ de Markov (MRF) (Besag 1974). Un champ aléatoire est un champ de Markov si la probabilité conditionnelle d'une variable u_s connaissant toutes les autres variables $\mathbf{u}_{\mathcal{S} \setminus s}$ ne dépend que des variables voisines :

$$p(u_s | \mathbf{u}_{\mathcal{S} \setminus s}) = p(u_s | \mathbf{u}_{\partial s}) \quad (2.2)$$

où ∂s désigne le voisinage du site s : $\partial s = \{t \in \mathcal{V} : (s, t) \in \mathcal{A}\}$.

Nous pouvons aussi écrire l'énergie sous une forme paramétrique : $E(\mathbf{u}) = \sum_{k=1}^p \theta_k E_k(\mathbf{u}) = \boldsymbol{\theta}^T \underline{E}(\mathbf{u})$ où $\boldsymbol{\theta} \in \mathbb{R}^p$ est un vecteur de paramètres contrôlant l'intensité des interactions du modèle et $\underline{E}(\cdot)$, défini sur les cliques du système de voisinage, est une fonction qui prend ses valeurs dans \mathbb{R}^p . Dans ces conditions, nous obtenons

$$p(\mathbf{u} | \boldsymbol{\theta}) = \frac{1}{Z(\boldsymbol{\theta})} \exp\left(-\boldsymbol{\theta}^T \underline{E}(\mathbf{u})\right) \quad (2.3)$$

où la constante de normalisation $Z(\boldsymbol{\theta})$ dépend du vecteur de paramètres $\boldsymbol{\theta}$:

$$Z(\boldsymbol{\theta}) = \sum_{\mathbf{u} \in \mathcal{U}} \exp\left(-\boldsymbol{\theta}^T \underline{E}(\mathbf{u})\right) < +\infty. \quad (2.4)$$

Une des difficultés rencontrées dans l'usage de ces modèles de Gibbs est liée à l'impossibilité pratique de calculer la constante de normalisation $Z(\boldsymbol{\theta})$. Plusieurs approximations ont ainsi été proposées reposant par exemple sur les notions de pseudo-vraisemblance (Besag 1975). Nous rencontrons le même type de difficultés lorsqu'il s'agit de manipuler les probabilités *a posteriori* associées à ces modélisations par champs de Markov. Néanmoins, le paramètre $\boldsymbol{\theta}$ qui maximise la log-vraisemblance au vu d'une configuration observée \mathbf{u} est solution de l'équation

$$\underline{E}(\mathbf{u}) = \mathbb{E}_{\boldsymbol{\theta}}[\underline{E}(\mathbf{U})] \quad (2.5)$$

où \mathbf{U} est un champ aléatoire et $\mathbb{E}_{\theta}[\cdot]$ désigne l'espérance relativement à $p(\cdot|\theta)$. On vérifie aisément que la fonction log-vraisemblance est strictement concave et admet donc une solution unique. En pratique, une descente de gradient sur la fonction de vraisemblance permet bien souvent d'estimer le paramètre θ . Des techniques itératives du type Monte-Carlo sont généralement nécessaires pour traiter des modèles plus riches. Pour des structures de voisinage plus complexes mais toujours possibles, l'inférence sur le paramètre θ est d'autant plus difficile que le calcul de la constante de normalisation $Z(\theta)$ devient insoluble.

2.1.2 Approche bayésienne et sélection de modèles

Dans la perspective de mettre en compétition M champs de Gibbs pour décrire une image donnée \mathbf{u} , nous associons à chaque modèle une fonction de vraisemblance

$$p(\mathbf{u}|\theta_m) = \frac{1}{Z(\theta_m)} \exp\left(-\theta_m^T \underline{E}_m(\mathbf{u})\right) \quad (2.6)$$

où $\theta_m \in \Theta_m$ et $Z(\theta_m)$ désigne la constante de normalisation. Typiquement, il peut s'agir de mettre en compétition M structures de voisinage d'ordres différents $s \stackrel{m}{\sim} t$ ($0 \leq m \leq M$). Dans le paradigme bayésien, nous avons recours au calcul des probabilités *a posteriori* de ces modèles. En particulier, θ_m est considéré comme une quantité aléatoire de densité $p(\theta_m)$, dite loi *a priori* sur θ_m . L'estimation bayésienne est basée sur le calcul de la loi *a posteriori* de θ_m , c'est-à-dire la loi conditionnelle de θ_m sachant \mathbf{u} dont la densité, notée $p(\theta_m|\mathbf{u})$, est définie comme

$$p(\theta_m|\mathbf{u}) \propto p(\mathbf{u}|\theta_m)p(\theta_m) \quad (2.7)$$

d'après le théorème de Bayes. Considérons à présent M modèles bayésiens paramétriques et munissons l'espace des modèles d'une loi de probabilité *a priori*. Un choix de modèle bayésien est basé sur la probabilité *a posteriori* des différents modèles

$$\begin{aligned} p(\mathcal{M} = m|\mathbf{u}) &\propto p(\mathcal{M} = m) \int_{\Theta_m} p(\mathbf{u}, \theta_m|\mathcal{M} = m) d\theta_m \\ &\propto p(\mathcal{M} = m) \int_{\Theta_m} p(\mathbf{u}|\theta_m)p(\theta_m) d\theta_m, \end{aligned} \quad (2.8)$$

c'est-à-dire la marginale de la distribution *a posteriori* sur $(\mathcal{M}, \theta_0, \dots, \theta_{M-1})$ connaissant \mathbf{u} . On choisira le modèle m' si $p(\mathcal{M} = m'|\mathbf{u}) \geq p(\mathcal{M} = m|\mathbf{u}), \forall m \neq m'$.

Mettons maintenant en lumière quelques difficultés associées à cette méthodologie lorsqu'elle est mise œuvre en analyse d'images :

- La distribution *a posteriori* des modèles est très sensible au choix des lois *a priori* $p(\theta_m)$ des paramètres des modèles. Il est par ailleurs important que ces lois soient équitables, ce qui n'est pas simple à obtenir.
- Les distributions *a priori* doivent correspondre à des densités d'intégrale finie afin de pouvoir mettre en œuvre la méthodologie bayésienne de choix de modèles.
- Pour des modèles complexes, nous ne pouvons pas toujours calculer explicitement (2.9). Des méthodes d'approximation stochastique sont donc nécessaires mais également lourdes

à mettre en œuvre.

- Enfin, lorsque le nombre de modèles en compétition augmente, il n'est pas toujours possible de calculer explicitement la loi *a posteriori* des modèles.

Toutes ces difficultés ont justifié une recherche intensive en traitement d'images. Au final, la sélection de modèle est rarement mise en œuvre et un seul modèle de Markov m_0 est choisi arbitrairement, mais avec bon sens. L'effort de recherche s'est alors davantage orienté vers la résolution de « problèmes mal posés » au sens d'Hadamard (Geman & Geman 1984, Geman, Geman, Graffigne & Dong 1990, Winkler 1995, Chalmond 2000, Hurn, Husby & Rue 2003, Descombes 2004) et la minimisation globale des énergies sous-jacentes (Greig et al. 1989, Boykov et al. 2001, Ishikawa 2003). La segmentation d'images figure parmi les problèmes fondamentaux à résoudre depuis très longtemps puisqu'il se pose en préalable à la plupart des schémas d'interprétation ou de décision.

2.1.3 Segmentation bayésienne d'images

L'analyse bayésienne d'images procure des solutions élégantes aux problèmes de segmentation. Elle a été proposée à l'origine par les frères Geman (1984) pour résoudre des problèmes de restauration avec préservation des discontinuités spatiales, puis de segmentation d'images (Geman et al. 1990). Elles mettent en jeu des variables discrètes cachées et observées. L'approche retenue, toujours d'actualité, nécessite la définition d'une connaissance *a priori* $p(\mathbf{u})$ sur le champ caché \mathbf{u} , généralement donnée par la définition d'un champ de Markov (Winkler 1995, Pérez 1998, Chalmond 2000, Hurn et al. 2003, Descombes 2004). On considère un second champ aléatoire $\mathbf{f} = (f_1, \dots, f_N)$ qui lui sera observé, mais dont l'espace d'états est souvent plus grand (il y a coïncidence des espaces d'états en restauration ou déconvolution d'images). Par la suite, on supposera que les supports des deux champs aléatoires se correspondent pour simplifier l'exposé. Le lien entre les variables observées et les variables cachées est obtenu au travers la spécification d'une fonction de vraisemblance $p(\mathbf{f}|\mathbf{u})$. Cette fonction de vraisemblance s'appuie bien souvent sur le modèle physique de formation de l'image et peut se factoriser avec des hypothèses, parfois un peu fortes, d'indépendance conditionnelle, c'est-à-dire sous la forme suivante :

$$p(\mathbf{f}|\mathbf{u}) \propto \prod_s p(f_s|\mathbf{u}_{\partial s}) \quad (2.9)$$

où ∂s désigne un ensemble restreint de sites voisins du site s . Par la suite, le problème posé est celui de l'estimation du champ caché $\mathbf{u} = (u_1, \dots, u_N)$ à partir du champ observé $\mathbf{f} = (f_1, \dots, f_N)$, c'est-à-dire l'inférence de \mathbf{u} une fois la distribution *a posteriori* $p(\mathbf{u}|\mathbf{f})$ établie. Plusieurs estimateurs minimisant un risque bayésien conditionnel ont été étudiés (marginales *a posteriori* (MPM) (Marroquin & Poggio 1987), Champ Moyen (MF) (Celeux, Forbes & Peyrard 2003)) mais la plupart nécessite de mettre en œuvre des simulations de Monte-Carlo, et de calculer des marginales locales de manière empirique et coûteuses, dans la mesure où le nombre de variables à estimer est relativement grand. Au prix de l'abandon de certaines qualités de ces estimateurs (mesure de confiance sur l'estimation), l'essentiel des travaux en analyse d'images a porté sur l'estimateur du maximum *a posteriori* qui vise surtout à extraire le mode de la

distribution $p(\mathbf{u}|\mathbf{f})$:

$$\hat{\mathbf{u}} = \arg \max_{\mathbf{u}} p(\mathbf{u}|\mathbf{f}) = \arg \max_{\mathbf{u}} p(\mathbf{f}|\mathbf{u})p(\mathbf{u}). \quad (2.10)$$

Le terme a priori $p(\mathbf{u})$ permet d'évaluer la propension de \mathbf{u} à se réaliser dans l'absolu, indépendamment de l'observation \mathbf{f} . La marginale $p(\mathbf{f})$ est non informative au regard de \mathbf{u} . Il s'agit ici de favoriser certains modèles, soit parce qu'ils sont objectivement plus rencontrés, soit parce qu'ils donnent des interprétations plus réalistes ou raisonnables dans le cadre d'une application donnée. Même si les champs de Markov sont justement conçus pour spécifier cet *a priori*, d'autres outils probabilistes sont également possibles : critère d'information (Leclerc 1989), modèle statistique et paramétrique de forme (Grenander, Chow & Keenan 1991, Grenander & Miller 1994, Baddeley & Van Lieshout 1993, Descombes 2004).

L'exemple du modèle d'images constantes par morceaux

Pour préciser l'idée précédente, nous considérons à présent quelques modèles de segmentation d'images établis dans un cadre bayésien, dans le cas particulier où le lien entre les variables cachées et les variables observées est issu du modèle d'images suivant :

$$f_s = a_{u_s} + \varepsilon_s, \quad s \in \{1, \dots, N\} \quad (2.11)$$

où $u_s = \{1, \dots, L\}$ est un numéro de classe, ε_s représente un bruit blanc gaussien de moyenne nulle et de variance $\sigma_{u_s}^2$: $\varepsilon_s \sim \mathcal{N}(0, \sigma_{u_s}^2)$. Les paramètres des classes (a_ℓ, σ_ℓ^2) , $\ell = \{1, \dots, L\}$ sont supposés estimés par avance ou sont des inconnues du problème, au même titre que le champ caché \mathbf{u} . Si on souhaite décomposer l'image en L classes, la fonction de vraisemblance s'exprime simplement de la manière suivante :

$$p(\mathbf{f}|\mathbf{u}) = \exp -\frac{1}{2} \left(\sum_{\ell=1}^L \sum_{s:u_s=\ell} \left(\frac{(f_s - a_\ell)^2}{\sigma_\ell^2} + \log(2\pi\sigma_\ell^2) \right) \right).$$

Pour traiter le cas des images texturées, une démarche usuelle consiste à calculer des attributs de texture à l'aide de bancs de filtres (Geman et al. 1990, Zhu & Mumford 1997) et d'élaborer une fonction de vraisemblance telle que nous l'avions également définie dans (Kervrann & Heitz 1995). Le modèle *a posteriori* à considérer se décline selon le choix du terme *a priori*. Celui-ci favorise l'émergence de régions homogènes délimitées par des bords réguliers et d'un faible nombre de régions (Geman & Geman 1984, Blake & Zisserman 1987, Geman et al. 1990) ou d'objets de forme géométrique (Grenander et al. 1991, Grenander & Miller 1994, Baddeley & Van Lieshout 1993, Descombes 2004) ou textuelle (Zhu & Mumford 1997) connus :

- Dans le cadre de la modélisation d'images par champs de Markov, le modèle de Potts (modèle d'Ising étendu à L classes) est bien souvent préconisé pour un espace d'états fini (modèle de Potts isotrope), ce qui conduit à

$$p(\mathbf{u}|\mathbf{f}) \propto \exp \left(\theta \sum_{s \sim t} \mathbb{1}[u_s = u_t] \right) p(\mathbf{f}|\mathbf{u})$$

où $\mathbb{1}[\cdot]$ est la fonction indicatrice et $\theta > 0$ est un scalaire qui encourage deux variables voisines à être identiques. La somme $\sum_{s \sim t}$ signifie que la sommation est effectuée sur l'ensemble des paires de sites voisins. Le modèle de Potts a tendance à isoler un petit nombre de régions compactes, c'est-à-dire dont le périmètre est aussi faible que possible, un *a priori* qui sera également repris dans le cadre de la modélisation variationnelle (Lellmann, Becker & Schnörr 2009, Pock, Chambolle, Cremers & Bischof 2009).

- Plus récemment, l'idée de spécifier directement la loi *a posteriori* a suscité un certain engouement (Kumar, Torr & Zisserman 2005), laissant une plus grande liberté pour modéliser les liens souvent complexes qui relient les variables cachées aux variables observées, et une certaine souplesse sur la notion de contexte spatial. On considère alors la distribution *a posteriori* suivante :

$$p(\mathbf{u}|\mathbf{f}) \propto \exp - \left(\sum_{s \sim t} \phi_1(u_s, u_t) + \sum_{s \in \mathcal{S}} \sum_{s' \in V_s} \phi_2(u_s, u_{s'}, \mathbf{f}) - \sum_{s \in \mathcal{S}} \log p(\mathbf{f}|u_s) \right)$$

où V_s désigne un voisinage du site s qui peut être différent de ∂s . Dans le cas particulier que nous considérons, nous avons

$$\begin{cases} \phi_1(u_s, u_t) &= -\theta_1 \mathbb{1}[u_s = u_t], \\ \phi_2(u_s, u_{s'}, \mathbf{f}) &= -\theta_2 \mathbb{1}[u_s = u_{s'}] \frac{e^{-\alpha(f_s - f_{s'})^2}}{\text{dist}(s, s')}. \end{cases}$$

Dans cette formulation de champs aléatoires conditionnels (CRF) (Kumar & Hebert 2006, Kohli & Torr 2006, Szummer, Kohli & Hoiem 2008), le potentiel $\phi_2(\cdot)$, contrôlé par les paramètres θ_2 et α , fait notamment intervenir les variables observées (Kumar et al. 2005). Une telle formulation est difficile à obtenir en adoptant une décomposition bayésienne classique.

- Suite aux travaux du groupe de Grenander à l'université de Brown (Grenander et al. 1991, Grenander & Miller 1994) sur les modèles génératifs de formes, des approches de segmentation dédiées à l'extraction de formes d'objets connus *a priori* ont connu un essor considérable ces vingt dernières années (Baddeley & Van Lieshout 1993, Descombes 2004). L'idée est de considérer des graphes orientés et un groupe de transformations dans un cadre markovien. Si on note $\mathbf{v} = (v_0, \dots, v_{K-1})$, les noeuds du graphe reliés entre eux par des arêtes, nous pouvons définir une approximation polygonale de la silhouette des objets à extraire. En raison des propriétés markoviennes d'indépendance conditionnelle sous-jacentes, des formes complexes et probabilisées sont finalement obtenues en appliquant le groupe de transformations à chaque noeud du graphe. Par la suite, il est toujours possible d'établir une correspondance entre la loi *a priori* $p(\mathbf{v})$ sur la forme polygonale et le champ caché \mathbf{u} . On peut noter ainsi $u_s = 1$ lorsque le site s est à l'intérieur de la forme polygonale et $u_s = 0$ si le site appartient à l'arrière plan pour lequel nous n'avons pas de connaissance *a priori* (Grenander et al. 1991, Chesnaud, Réfrégier & Boulet 2000). L'avantage procuré par ces modélisations de plus haut niveau que les modèles markoviens établis sur une grille régulière de pixels (Geman & Geman 1984), est de pouvoir générer des échantillons de la loi cible qui sont des représentations fidèles des objets observés (Zhu & Mumford 1997, Zhu 1999, Hurn et al. 2003). Enfin, l'analyse simultanée de plusieurs objets est toujours possible mais implique une com-

plexité croissante et la nécessité de gérer les occlusions entre les différentes formes polygonales par des algorithmes d’optimisation stochastiques (Grenander & Miller 1994, Hurn et al. 2003, Descombes 2004).

Un apprentissage statistique préalable des solutions minimisantes à partir d’une population représentative permet également de contraindre *a priori* l’espace des solutions admissibles (Cootes, Taylor, Cooper & Graham 1995, Cremers, Weickert & Schnörr 2002, Cremers, Kohlberger & Schnörr 2003). L’approche développée au cours de ma thèse reposait justement sur une modélisation paramétrique très compacte d’une base de contours (Kervrann & Heitz 1998, Kervrann 1998, Kervrann & Heitz 1999) associés à des vues particulières d’une classe d’objets et de déformations de la famille considérée. D’un point de vue statistique, cela se résume en partie à une décomposition modale (Analyse en Composantes Principales) des déformations observées sur une base de contours ou de textures, et à la construction d’une densité de probabilité sur la famille d’objets à extraire ou à localiser. On peut considérer enfin les travaux de Tu & Zhu (2002) comme un aboutissement de cette démarche de modélisation stochastique, mettant en évidence leurs performances indéniables mais aussi leurs difficultés de mise en œuvre.

Un exemple d’application en cryo-tomographie électronique

Nous avons tout récemment exploré la modélisation par champs aléatoires conditionnels (CRF) (Lafferty, McCallum & Pereira 2001, Kumar & Hebert 2006) pour analyser des images de cryo-tomographie électronique (Blestel, Kervrann & Chrétien 2010a), une technologie d’imagerie qui permet d’étudier des échantillons biologiques dans leur état natif à l’échelle nanométrique. L’objectif de ce projet débuté en 2008 dans le cadre de la thèse de Sophie Blestel, sous ma direction et celle de Denis Chrétien (UMR 6026, Rennes), est de mettre au point des algorithmes originaux permettant d’extraire des assemblages macromoléculaires (microtubules polymérisés et assemblés *in vitro*, centrosomes et axonèmes purifiés, anneaux de tubuline...). Dans cette contribution récente (en soumission) (Blestel et al. 2010a), le problème de segmentation est formulé comme un problème d’estimation au sens du maximum *a posteriori* de variables cachées $\mathbf{U} = \{U_s, s \in \mathcal{S}\}$ et $U_s \in \{F, O\}$, à partir du champ observé \mathbf{f} . On considère ainsi qu’un pixel de l’image appartient soit au fond (‘F’) soit à la classe d’objets d’intérêt (‘O’). Dans le cadre des Champs Aléatoires Conditionnels, nous modélisons directement la loi *a posteriori*, ce qui permet immédiatement de poser le problème d’estimation comme un problème de minimisation d’une énergie globale $E(\mathbf{u}, \mathbf{f})$:

$$E(\mathbf{u}, \mathbf{f}) = \sum_{s \in \mathcal{S}} D(u_s, \mathbf{f}) + \sum_{s \sim t} I(u_s, u_t, \mathbf{f}) \quad (2.12)$$

où $D(\cdot)$ est un terme de fidélité et $I(\cdot)$ est un potentiel d’interaction permettant d’imposer une certaine régularité à la carte de segmentation. Dans notre cas particulier nous avons considéré un modèle d’Ising $I(u_s, u_t, \mathbf{f}) = -\theta \mathbb{1}[u_s = u_t], \theta > 0$, et un terme de fidélité qui prend en compte le contexte spatial décrit par des motifs locaux en chaque site de l’image. On caractérise chacune des deux classes par une collection de motifs (carrés de taille $\sqrt{n} \times \sqrt{n}$) assez représentatifs des deux populations d’intérêt et extraits de l’image à traiter. Le terme de fidélité s’exprime comme

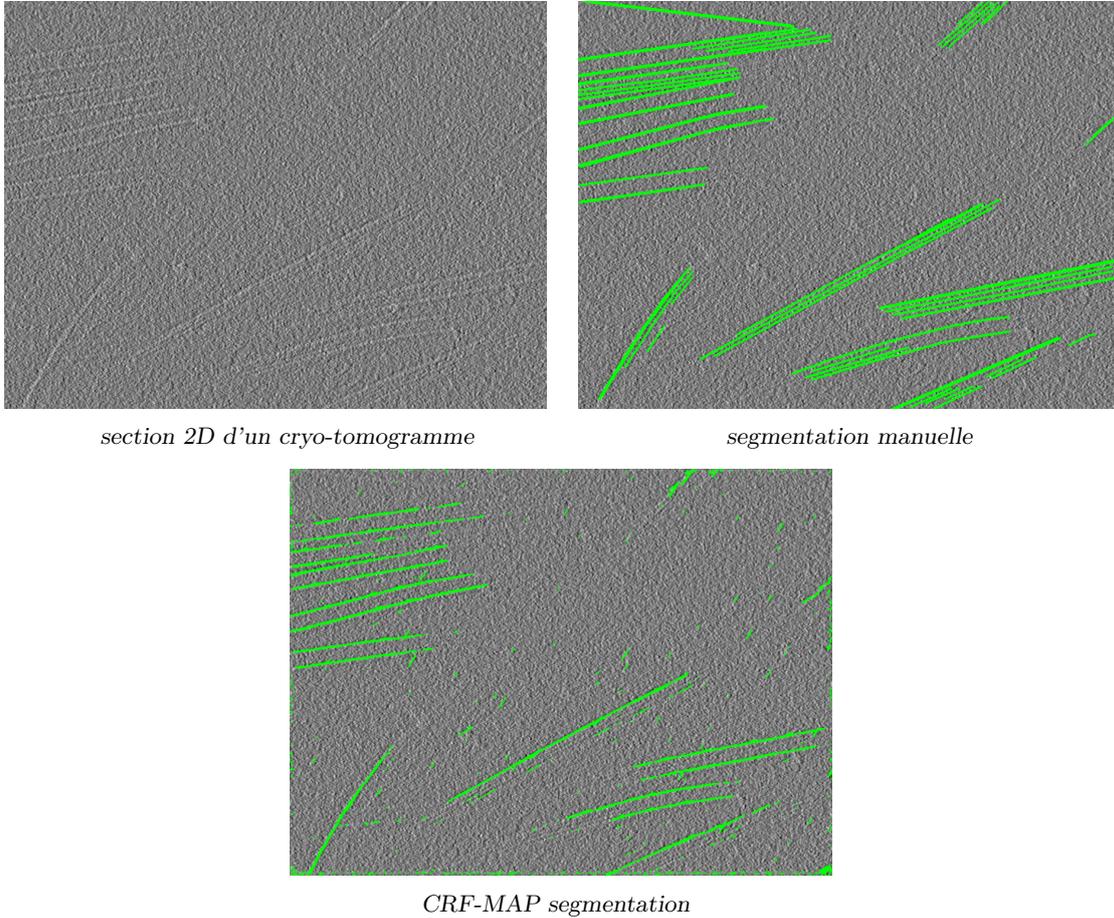


FIG. 2.1. Segmentation d'une section 2D extraite d'un cryo-tomogramme 3D.

la distance euclidienne entre le motif observé et les motifs représentatifs des deux classes extraits de l'image \mathbf{f} :

$$D(u_s, \mathbf{f}) = \|\mathbf{z}_s(\mathbf{f}) - \mathbf{z}_s(u_s, \mathbf{f})\|^2 \quad (2.13)$$

où $\mathbf{z}_s(\mathbf{f})$ est un vecteur de dimension n représentant le motif observé et centré au pixel s et, $\mathbf{z}_s(u_s, \mathbf{f})$ est le « meilleur » motif qui représente la classe u_s au pixel s . Une étape d'apprentissage est nécessaire pour élaborer un dictionnaire de motifs représentatifs de la classe d'objets d'intérêt. L'énergie globale est finalement minimisée par un algorithme « *min cut/max flow* » (Greig et al. 1989, Boykov et al. 2001) qui permet de déterminer très rapidement le minimum global de l'énergie $E(\mathbf{u}, \mathbf{f})$. Quelques résultats sont présentés sur la figure 2.1 illustrant le potentiel de l'algorithme contrôlé par θ et la taille n des motifs (Blestel et al. 2010a). L'extension au traitement de volumes a également été traité dans le cadre de cette étude (Fig. 2.2).

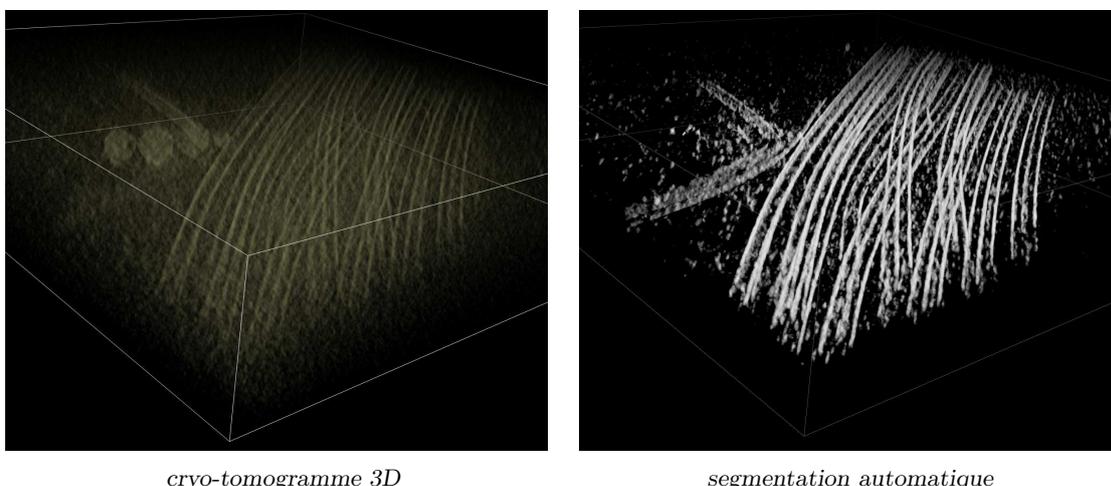


FIG. 2.2. CRF-MAP segmentation d'un cryo-tomogramme 3D.

Conclusion

La modélisation statistique des images à partir d'éléments picturaux locaux (i.e. motifs) (Lee et al. 2003, Koloydenko & Geman 2006), éventuellement dans le cadre formel des champs aléatoires conditionnels, s'inscrit dans la continuité des travaux initiés dans (Geman et al. 1990, Freeman, Pasztor & Carmichael 2000, Roth & Black 2005) sur la modélisation markovienne des images (Kervrann & Heitz 1995). Ce travail a été poursuivi dans le cadre du débruitage d'images et de la détection de changements dans une paire d'images décrits dans le chapitre 3.

D'une manière générale, les modèles exposés impliquent plusieurs termes d'énergie contrôlés par des paramètres qu'il s'agit également d'estimer par inférence bayésienne ou par des techniques de calibration (Azencott 1987, Chalmond 2000, Younes 2000). Les statisticiens se sont attachés à traiter cette question au prix d'une complexité algorithmique nécessitant l'usage de techniques coûteuses de Monte-Carlo. La communauté de vision par ordinateur et de traitement d'images a davantage porté son attention sur la minimisation globale des fonctionnelles d'énergie associées à ces modèles bayésiens, avec un succès certain.

2.2 Modèles variationnels

L'estimateur MAP (2.10), très populaire car relativement plus simple à mettre en œuvre que les estimateurs bayésiens MPM ou MF, cherche à calculer le mode de la distribution $p(\mathbf{u}|\mathbf{f})$, ce qui revient à minimiser l'énergie suivante :

$$E(\mathbf{u}, \mathbf{f}) = -\log p(\mathbf{u}) - \log p(\mathbf{f}|\mathbf{u}). \quad (2.14)$$

Il est donc très possible d'interpréter les termes d'énergie issus d'une formulation variationnelle et déterministe, dans un cadre probabiliste (Mumford (1994)). Dans cette partie de chapitre, nous décrivons deux modèles variationnels qui ont suscité un intérêt très important ces vingt dernières années de la part des communautés de vision par ordinateur, de traitement d'images et des mathématiques appliquées. Afin de clarifier les notations dans un cadre continu désormais, l'image observée est une fonction $f : \Omega \rightarrow \mathbb{R}$ (image de niveaux de gris, définie sur un domaine $\Omega \subset \mathbb{R}^2$) et l'inconnue du problème est également une fonction $u : \Omega \rightarrow \mathbb{R}$. Par la suite, $u(x)$ désigne la valeur de u au pixel $x \in \Omega$.

2.2.1 Minimisation de la variation totale

Un problème très commun dans bon nombre de domaines applicatifs est la résolution du problème de la déconvolution, un « problème inverse » notoirement connu comme difficile en traitement du signal et de l'image. Connaissant un modèle de dégradation qui introduit une flou sur l'image cherchée, on vise à reconstruire une image u à partir du modèle d'observation suivant :

$$f = Hu + \varepsilon \quad (2.15)$$

où $Hu = h \star u$ représente la convolution de u par un noyau h (fonction de transfert connue ou non) et ε désigne un terme de bruit, bien souvent assimilé à un bruit gaussien de moyenne nulle et de variance σ^2 . L'« inversion » directe de ce modèle conduit à des solutions peu satisfaisantes (oscillations importantes) à cause du terme de bruit amplifié lors de la reconstruction. Il est donc approprié d'imposer des contraintes *a priori* sur la solution recherchée. La première idée, issue des travaux en traitement du signal, consistait à choisir une régularisation du type Tikhonov (Tikhonov & Arsenin 1977) qui impose une solution régulière en tout point de l'image. Si on se place désormais dans un cadre de restauration ($Hu = u$), ceci revient à sélectionner, parmi toutes les fonctions u , la solution qui minimise

$$\int_{\Omega} |Du|^2 \text{ sous les contraintes } \int_{\Omega} u = \int_{\Omega} f \text{ et } \|u - f\|^2 = \sigma^2 \quad (2.16)$$

où Du représente le gradient spatial de u . Néanmoins, les images réelles contiennent des discontinuités très nettes qui ne sont pas cohérentes avec le modèle de Tikhonov. La solution du problème précédent, qui revient simplement à appliquer un filtre de Wiener dans le domaine de Fourier, mais produit en effet des images exagérément floues sur les zones de contours. Pour contourner cette difficulté, Rudin et al. (1992) ont préconisé au début des années 90 le modèle de « variation totale » en substituant la norme L_1 à la norme L_2 dans le terme de régularité. Cette suggestion est d'autant plus justifiée qu'il est assez aisé de constater expérimentalement que la distribution du module du gradient dans la plupart des images naturelles, suit une loi de Laplace (exponentielle). D'après (2.14), le modèle (2.16) favorise donc l'émergence de zones homogènes tout en restaurant les discontinuités de l'image. De manière formelle, l'espace des solutions à considérer est celui des fonctions $u \in L^1$ pour lesquelles les variations $|Du|$ (au sens des distributions) sont

bornées (Rudin et al. 1992). Finalement, on résout le problème de minimisation suivant :

$$\text{minimiser } \int_{\Omega} |Du| + \frac{\beta}{2} \int_{\Omega} (u - f)^2 \quad (2.17)$$

où β s'interprète comme un multiplicateur de Lagrange et la seconde contrainte est apparentée à un terme de fidélité, très semblable à une log-vraisemblance. La solution est obtenue par résolution d'une équation aux dérivées partielles dépendant d'une variable $t > 0$ analogue à une variable temporelle et issue d'une dérivation selon Euler-Lagrange de la fonctionnelle d'énergie. Soit encore

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{Du}{|Du|} \right) - \beta(u - f) \quad \text{avec } u(x, 0) = f(x), \quad x \in \Omega \quad (2.18)$$

et des contraintes aux bords. Notons que plusieurs algorithmes ont été proposés ces dernières années pour déterminer u , mais étrangement, ils peuvent conduire à des solutions visuellement assez différentes (Rudin et al. 1992, Dibos & Koepfler 1998, Chambolle 2004). Notons que le cadre statistique de l'estimation robuste (Black & Rangarajan 1996a, Black, Sapiro, Marimont & Heeger 1998a), très courant, permet d'établir un lien entre le modèle de Rudin et al. (1992) et la modélisation statistique des images. Pour cela, il suffit de considérer le cas général suivant

$$\text{minimiser } \int_{\Omega} \Phi(|Du|) + \frac{\beta}{2} \int_{\Omega} (u - f)^2 \quad (2.19)$$

où $\Phi(\cdot)$ une fonction positive croissante définie sur \mathbb{R} . Cette énergie coïncide avec (2.17) si $\Phi(z) = z$. Très récemment, une nouvelle interprétation bayésienne (Espérance *a posteriori*) du modèle de Rudin et al. (1992), a été proposée par Louchet & Moisan (2008) dans un cadre continu, :

$$\hat{u}_{LSE} = \frac{\int_{\mathbb{R}^{\Omega}} u e^{-\frac{1}{2\sigma^2} (\int_{\Omega} (u-f)^2 + \lambda_{TV} \int_{\Omega} |Du|)} du}{\int_{\mathbb{R}^{\Omega}} e^{-\frac{1}{2\sigma^2} (\int_{\Omega} (u-f)^2 + \lambda_{TV} \int_{\Omega} |Du|)} du} \quad (2.20)$$

où \hat{u}_{LSE} désigne l'estimateur au sens de l'« erreur quadratique moyenne » et λ_{TV} est une constante positive. Des approximations par des techniques stochastiques de Monte-Carlo sont néanmoins nécessaires pour s'affranchir du calcul de la constante de normalisation, si problématique en analyse bayésienne d'images.

Si le modèle de Rudin et al. (1992) a aussi connu un tel succès, c'est probablement grâce aux propriétés géométriques remarquables associées à l'évolution des lignes de niveau de l'image. En effet, minimiser $\int_{\Omega} |Du|$ sans contrainte revient à résoudre une équation aux dérivées partielles du type

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{Du}{|Du|} \right) = \operatorname{curv}(u). \quad (2.21)$$

où $\operatorname{div}(\cdot)$ désigne l'opérateur divergence (Alvarez et al. 1993, Malladi & Sethian 1996). Sur le plan numérique, on itère cette équation qui déplace les lignes de niveau selon leur normale et

proportionnellement à leur courbure (notée $\text{curv}(\cdot)$). Asymptotiquement, la solution est triviale et correspond à une image constante. Néanmoins le caractère transitoire voire l'évolution vers la solution optimale est plus intéressant à étudier. Le nombre d'itérations pour améliorer l'image est néanmoins aussi problématique que le choix du paramètre β . Le terme de fidélité aux données permet en effet de détecter automatiquement la solution la plus pertinente pour approcher l'image originale. Plus fondamentalement, la modification de cette équation d'évolution a permis d'établir le lien avec un grand nombre de filtres non-linéaires (e.g. filtre médian) étudiés en morphologie mathématique dans un cadre discret (Salembier & Serra 1995). On considère alors les équations aux dérivées partielles de la forme suivante (mouvement par courbure moyenne) :

$$\frac{\partial u}{\partial t} = |Du|F(\text{curv}(u)) \quad (2.22)$$

où F est une fonction croissante. L'approche axiomatique, que je ne décrirai pas dans ce document, et très répandue dans la communauté « espace-échelle » (« *scale-space* »), a permis d'étudier les notions de causalité et d'invariance aux changements de contraste. Les notions d'existence, d'unicité et de stabilité de la solution pour une condition initiale donnée, ont été également établies pour des équations aux dérivées partielles de ce type (Alvarez et al. 1992, Catte, Lions, Morel & T. 1992, Alvarez et al. 1993).

Pour conclure, notons enfin que minimiser $\int_{\Omega} \Phi(|Du|)$ revient à opérer une diffusion linéaire parfaitement isotrope si $\Phi(z) = z^2$ ou une diffusion non-linéaire si $\Phi(z) = 1 - e^{-z^2/K}$ (Perona & Malik 1990a). D'autres formulations ont été proposées dans le cadre de la diffusion non-linéaire anisotrope pour simplifier les images. Une formulation assez générique consiste à imposer une diffusion selon une direction imposée par le tenseur de structure $\mathbf{T} = G_{\sigma} \star (DuDu^T)$. On considère alors des équations d'évolution du type

$$\frac{\partial u}{\partial t} = \text{div}(\mathbf{T}Du) \quad (2.23)$$

où G_{σ} est un noyau gaussien de convolution (Weickert 1999a, Tschumperlé 2006).

2.2.2 Segmentation d'images et équations aux dérivées partielles

Tout aussi populaires que les modèles variationnels exposés précédemment, les modèles variationnels de segmentation d'images ayant pour origine les travaux de Kass et al. (1987), ont suscité des travaux nombreux et incessants pour des domaines applicatifs aussi variés que l'imagerie médicale, l'imagerie fluide, la télédétection... Malgré la pertinence des formulations proposées (Zhu & Yuille 1996, Yezzi, Tsai & Willsky 1999, Chan & Vese 2001, Cremers et al. 2002, Heiler & Schnörr 2005), mais aussi dans un cadre bayésien (Mumford 1994, Zhu & Yuille 1996, Brox & Cremers 2009) voire d'estimation non-paramétrique (Brox & Cremers 2009), les critères globaux de segmentation d'images ont pour origine la fonctionnelle de Mumford & Shah (1989) :

$$E(u, B) = \int_{\Omega \setminus B} |Du|^2 + \beta \int_{\Omega \setminus B} (u - f)^2 + \nu \mathcal{H}^1(B) \quad (2.24)$$

qui vise à chercher une image $u : \Omega \rightarrow \mathbb{R}$ à partir d'une image observée f , régulière en dehors d'un ensemble B (bords) où β est un paramètre positif, $\nu \geq 0$ est un paramètre d'échelle et \mathcal{H}^1 la mesure de Hausdorff unidimensionnelle (longueur de bords). Cette approche globale a le mérite de pouvoir isoler par simple minimisation, des régions régulières par morceaux séparées par des courbes fermées « rectifiables » (Morel & Solimini 1994). La mise en œuvre est néanmoins délicate même si des formulations implicites, autorisant des changements de topologie, ont sensiblement amélioré les schémas numériques (Osher & Sethian 1988, Caselles et al. 1995, Sethian 1996). Il n'en demeure pas moins qu'il faut toujours ajuster des paramètres ν et β en fonction de la taille, du nombre ou de la régularité des frontières des objets à segmenter dans l'image.

Un cas plus simple, mais finalement plus adapté au problème de la segmentation d'images, est la restriction aux régions constantes par morceaux (Blake & Zisserman 1987, Mumford & Shah 1989, Koepfler, Lopez & Morel 1994, Zhu & Yuille 1996, Schnörr 1998, Chan & Vese 2001, Brox & Cremers 2009) et qui correspond au cas limite $\beta \rightarrow \infty$. On considère alors la fonctionnelle d'énergie suivante :

$$E(f, D_1, \dots, D_P) = \sum_{p=1}^P \int_{D_p} (f - \bar{f}_p)^2 + \nu \sum_{p=1}^P \text{per}(D_p) \quad (2.25)$$

où $\text{per}(D_p)$ représente la longueur de périmètre de la région D_p , définissant ainsi une approximation dite « cartoon » (constante par morceaux) de l'image originale. L'image f est alors approchée par des fonctions constantes par morceaux dont les supports D_1, \dots, D_P sont séparés par des bords et \bar{f}_p est un scalaire qui représente l'intensité moyenne estimée sur le support D_p . Notons enfin qu'une relaxation convexe de cette énergie de segmentation a permis très récemment d'apporter un éclairage nouveau sur les solutions minimisantes (Chan, Esedoglu & Nikolova 2006, Lellmann et al. 2009, Pock et al. 2009). Pour cela, il a été proposé par exemple d'introduire des variables auxiliaires $\mathbf{w}(x) = (w_1(x), \dots, w_P(x))^T$ telles que $w_p(x) \in [0, 1]$ et $\sum_{p=1}^P w_p(x) = 1$. Le problème de la minimisation de l'énergie (2.26) revient alors à résoudre le problème d'optimisation suivant :

$$\min_{\mathbf{w}} \sum_{p=1}^P \int_{\Omega} w_p(x) (f(x) - \bar{f}_p)^2 dx + \nu_0 \sum_{p=1}^P \int_{\Omega} |Dw_p(x)| dx \quad (2.26)$$

où ν_0 est un paramètre d'échelle analogue à ν .

2.3 Un principe morphologique, statistique et variationnel pour la segmentation

Comme nous l'avons évoqué, l'analyse bayésienne d'images procure des solutions élégantes à ces problèmes de segmentation. Elle conduit en définitive à la formulation du problème comme celui de la minimisation globale d'une fonctionnelle d'énergie mêlant variables observables et paramètres du modèle. Cette fonctionnelle se décompose elle-même en deux termes distincts : un premier terme mesurant la qualité de l'approximation et un second terme qui encourage *a priori*

l'émergence de régions délimitées par des bords réguliers. Les modèles mis en jeu dans ce contexte peuvent être discrets (modèles markoviens (Geman & Geman 1984)) ou continus (contours actifs (Caselles et al. 1997), fonctionnelle de Mumford & Shah (1989) (Mumford & Shah 1989, Morel & Solimini 1994, Zhu & Yuille 1996, Chan & Vese 2001, Brox & Cremers 2009)). Cependant, la mise en œuvre de ces méthodes peut s'avérer délicate en raison de la charge calculatoire nécessaire pour mener la minimisation globale des énergies.

Nous nous intéressons ici aux questions de performance et de rapidité des schémas de segmentation non-supervisés s'appuyant sur une minimisation de fonctionnelle d'énergie. On peut remarquer qu'il n'est pas toujours aisé d'explicitier le minimum global de fonctions non-convexes utiles dans les applications (Morel & Solimini 1994). Il apparaît donc pertinent d'identifier et d'étudier les fonctionnelles pour lesquelles les solutions peuvent être déterminées à l'avance. On peut alors s'abstenir de recourir à des procédures itératives de minimisation lors de la mise en œuvre de l'algorithme de segmentation. Une alternative à cette démarche est de proposer des approximations convexes de ces fonctionnelles d'énergie comme nous venons de l'évoquer précédemment.

2.3.1 Représentation de l'image par ses lignes de niveau

Soit une image en niveaux de gris représentée par une fonction $u : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ semi-continue supérieurement où x désigne la position d'un pixel dans l'image. En pratique, une simple opération de seuillage sur u permet d'obtenir les ensembles de niveau :

$$L_\gamma(u) = \{x \in \mathbb{R}^2 : u(x) \geq \gamma\}. \quad (2.27)$$

Chaque ensemble de niveau L_γ contient plusieurs composantes connexes. Les bords topologiques de ces composantes connexes sont appelés lignes de niveau de u . Notons que ces lignes n'ont aucune propriété de régularité. Les frontières des composantes connexes des $L_\gamma(u)$ forment aussi la carte topographique de u (Monasse & Guichard 2000, Caselles, Coll & Morel 1999). Ces lignes de niveau sont globalement invariantes lorsqu'on applique un changement de contraste à u , i.e.

$$L_{g(\gamma)}(g \circ u) = L_\gamma(u) \quad (2.28)$$

où g est une fonction croissante de \mathbb{R} .

Cette représentation « morphologique » proposée par Matheron (1975), permet de traiter les ensembles de niveau séparément et de les combiner. Finalement, une reconstruction de l'image sans perte d'information est possible si on considère tous les ensembles de niveau :

$$u(x) = \sup\{\gamma : x \in L_\gamma(u)\}. \quad (2.29)$$

Pour obtenir une bonne représentation sur le plan numérique et respectueuse du théorème de Jordan, il est recommandé de calculer les ensembles de niveau « supérieurs » et « inférieurs » (Monasse & Guichard 2000) et d'améliorer la régularité des lignes de niveau à l'aide d'interpolations bilinéaires. Un filtrage linéaire ou non-linéaire peut être aussi préconisé pour lisser l'image au préalable en préservant si possible les structures géométriques, une démarche que nous adoptons par la suite.

2.3.2 Segmentation par minimisation de fonctionnelles d'énergie et sélection de lignes de niveau

Nous considérons ici le problème de la segmentation d'images lorsque l'image reconstruite est supposée constante par morceaux conformément au modèle simplifié de Mumford & Shah (1989). Soit une image observée $f : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ où $f(x)$ est l'intensité observée au pixel $x \in \Omega$. Soit $D_p \subset \Omega$, $p \in \{1, \dots, P\}$ une région ou un domaine de l'image. On suppose que la minimisation de l'énergie suivante (Kervrann, Hoebeke & Trubuil 2002, Kervrann & Trubuil 2002, Kervrann 2001, Kervrann, Hoebeke & Trubuil 1999, Kervrann, Hoebeke & Trubuil 2000) et inspirée de Beaulieu & Goldberg (1989), permet de partitionner l'image en régions (ou domaines) :

$$E(f, D_1, \dots, D_P) = \sum_{p=1}^P \int_{D_p} (f(x) - \bar{f}_p)^2 dx + \lambda \text{pen}(P, D_1, \dots, D_P) \quad (2.30)$$

où $\bar{f}_p = \frac{1}{|D_p|} \int_{D_p} f(x) dx$ désigne la luminance moyenne calculée sur le domaine D_p , $|D_p| = \int_{D_p} dx$, λ est un réel positif et $\text{pen}(P, D_1, \dots, D_P)$ est un terme *a priori* qui permet d'imposer des contraintes sur la partition.

Notre objectif est de sélectionner une collection \mathcal{C}_P de régions ($\{D_1, \dots, D_{P-1}\}$, $P \geq 1$) disjointes ($D_p \cap D_{p'} = \emptyset$, $p \neq p'$), strictement incluses dans Ω , simplement connexes, se superposant sur le fond D_P de l'image et telles que $\cup_{p=1}^P D_p = \Omega$. Si $P = 1$, l'image ne contient pas de régions. Par définition, D_P occupe un statut particulier puisque cette région n'est pas nécessairement connexe. Une segmentation optimale de l'image f (si elle existe) est par définition un minimum global de l'énergie :

$$(D_1^*, \dots, D_{P-1}^*) = \inf_{P \geq 1} \inf_{\{D_1, \dots, D_{P-1}\} \in \mathcal{C}_P} E(f, D_1, \dots, D_P). \quad (2.31)$$

Une minimisation directe de cette énergie sur les domaines inconnus des régions D_p et les paramètres \bar{f}_p s'avère extrêmement complexe.

Contribution principale en quelques mots

Pour certains choix de pénalisation, nous avons pu prouver que les frontières des régions qui minimisent le critère de segmentation correspondent à des lignes de niveau de f . Pour cela, nous devons prendre par exemple $\text{pen}(P, D_1, \dots, D_P) = P$ ou $\text{pen}(P, D_1, \dots, D_P) = \sum_{p=1}^{P-1} |D_p|$. Ce choix de fonctionnelle permet de pénaliser uniquement l'émergence d'un trop grand nombre de régions dans l'image et λ peut être interprété comme un paramètre d'échelle (Beaulieu & Goldberg 1989, Morel & Solimini 1994) voire comme un paramètre contrôlant la « parcimonie morphologique » de l'image. Le cas usuel $\text{pen}(P, D_1, \dots, D_P) = \sum_{p=1}^{P-1} \text{per}(D_p)$ qui revient à minimiser la longueur du périmètre de chaque région, ne peut pas être considéré dans notre cas. Ce choix particulier coïncide avec la fonctionnelle simplifiée (2.26) de Mumford & Shah (1989), repris avec succès par Chan & Vese (2001) il y a quelques années, et correspond enfin à une formulation continue du « modèle de Potts » évoqué en Section 2.1.3.

Description du minimiseur

Notre estimateur non-paramétrique \hat{u} est une image constante par morceaux et résulte de la minimisation globale de l'énergie (2.30). De manière équivalente, on cherche les domaines optimaux, c'est-à-dire

$$(\hat{D}_1, \dots, \hat{D}_P) = \arg \min_{P \geq 1} \arg \min_{\{D_1, \dots, D_{P-1}\} \in \mathcal{C}_P} E(f, D_1, \dots, D_P). \quad (2.32)$$

La question de l'existence d'un minimum global admissible pour des modèles de segmentation défini par Beaulieu & Goldberg (1989) ou Mumford & Shah (1989) a été abordée dans (Morel & Solimini 1994) mais ne sera pas discutée ici.

Sous les conditions d'existence d'un minimum global, nous avons proposé le lemme suivant pour une catégorie de fonctions de pénalisation que nous présentons par la suite (Kervrann et al. 2000, Kervrann & Trubuil 2002).

Lemme 1 *S'il existe un minimum global admissible dans le voisinage de la solution minimisante et qu'aucun minimum pathologique n'existe, alors l'ensemble des courbes qui minimisent l'énergie considérée est une sous-collection de lignes de niveau de f :*

$$f|_{\partial D_p} \equiv c_p, \quad p = 1, \dots, P-1.$$

En d'autres termes, le bord ∂D_p de chaque domaine D_p est une frontière d'un ensemble de niveau de la fonction f .

Nous avons prouvé ce lemme pour un domaine D et un fond D^c , où D^c désigne l'ensemble fermé complémentaire de D et où D_δ représente une faible perturbation de D , i.e. pour lequel la distance de Hausdorff $d_\infty(D_\delta, \Omega) \leq \delta$. Soit M un point fixe appartenant au bord ∂D et D_δ tel que $\partial D_\delta = \partial D$ excepté sur un petit voisinage de M . L'énergie admettant un minimum pour D , $f(M)$ doit être solution de l'équation, indépendamment du choix de M ,

$$A_0 + A_1 f(M) = 0 \quad (2.33)$$

où A_0 and A_1 sont définis comme

$$A_0 = \frac{m_1^2}{m_0^2} - \frac{(M_1 - m_1)^2}{(M_0 - m_0)^2}, \quad A_1 = -\frac{2m_1}{m_0} + \frac{2(M_1 - m_1)}{M_0 - m_0}$$

et $m_0 = \int_D dx$, $m_1 = \int_D f(x)dx$, $M_0 = \int_\Omega dx$ et $M_1 = \int_\Omega f(x)dx$. On déduit alors que $f(M)$ est constant lorsque M couvre ∂D .

Les premiers résultats associés à la mise en œuvre de l'algorithme correspondant et décrit par la suite, ont montré un gain sensible par rapport aux méthodes employées classiquement, souvent lourdes à implémenter. Nous nous sommes aussi posés la question de savoir si on ne pouvait pas proposer une famille d'énergie qui possède des minima globaux analogues en modifiant le terme de fidélité $E_d(f, D_1, \dots, D_P) = \sum_{p=1}^P \int_{D_p} (f(x) - \bar{f}_p)^2 dx$ ou bien, en considérant une classe plus

large de pénalisations. En fait, on peut choisir les modèles de contraste suivants :

$$E_d(f, D_1, \dots, D_P) = \int_{\Omega} f^2(x) dx - \sum_{p=1}^{P-1} \int_{D_p} f^2(x) dx, \quad (2.34)$$

$$E_d(f, D_1, \dots, D_P) = \sum_{p=1}^P \frac{1}{|D_p|} \int_{D_p} (f(x) - \bar{f}_p)^2 dx, \quad (2.35)$$

$$E_d(f, D_1, \dots, D_P) = - \sum_{p=1}^{P-1} (\bar{f}_p - \bar{f}_P)^2, \quad (2.36)$$

plus en rapport avec les besoins du praticien du traitement d'images, et ne reposant pas nécessairement sur un modèle de bruit gaussien, sous-jacent dans la modélisation originale. Le modèle (2.35) est approprié pour détecter des objets lumineux sur un fond sombre, ce qui est le cas pour des images de mesure de changements temporels. Le second modèle (2.36) vise à minimiser la variance de chacune des régions (Beaulieu & Goldberg 1989) tandis que le modèle (2.36) permet d'isoler les régions très contrastées par rapport au fond de l'image (O'Sullivan & Qian 1994, Yezzi et al. 1999).

Les modèles de régularité valides pour cette analyse variationnelle peuvent être rattachés *i)* à la classe des modèles markoviens sur des composantes connexes de l'image (Møller & Waagepetersen 1998) :

$$\text{pen}(P, D_1, \dots, D_P) = \sum_{p=1}^{P-1} \alpha_1 + \alpha_2 |D_p| + \alpha_3 |D_p|^2, \quad (2.37)$$

où les constantes α_1 , α_2 et α_3 sont des paramètres positifs ; *ii)* à l'entropie de l'image segmentée :

$$\text{pen}(P, D_1, \dots, D_P) = - \frac{1}{|\Omega|} \sum_{p=1}^P |D_p| \log(|D_p|/|\Omega|); \quad (2.38)$$

iii) à un modèle probabiliste (loi puissance) qui caractérise la distribution des aires des composantes connexes de l'image (Alvarez, Gousseau & Morel 1999) :

$$\text{pen}(P, D_1, \dots, D_P) = \sum_{p=1}^{P-1} (b \log(|D_p|) - a). \quad (2.39)$$

où les paramètres a et b sont obtenus par régression linéaire à partir de l'ensemble des aires des composante connexes qui décrivent l'image ou une classe donnée d'images. Le résultat principal que nous avons obtenu est que, avec ces choix restreints de modèles de contraste et de régularité, les courbes qui minimisent l'énergie de segmentation demeurent des lignes (ou surfaces) de niveau de l'image. Enfin, nous avons pu mettre en évidence que le nombre maximum d'objets extraits est directement lié au paramètre paramètre λ .

Lemme 2 *S'il existe une segmentation optimale correspondant au minimum de l'énergie*

$$\sum_{i=p}^P \int_{D_p} (f(x) - \bar{f}_p)^2 dx + \lambda \sum_{p=1}^{P-1} \alpha_1 + \alpha_2 |D_p| + \alpha_3 |D_p|^2 \quad (2.40)$$

et que $|D_p| > |D_{min}|, p = 1, \dots, P-1$, alors le nombre P^* d'objets extraits est borné :

$$P_{max} = \frac{1}{\lambda(\alpha_1 + \alpha_2 |D_{min}| + \alpha_3 |D_{min}|^2)} \int_{\Omega} \left(f(x) - \frac{1}{|\Omega|} \int_{\Omega} f(x) dx \right)^2 dx \quad (2.41)$$

Implémentation numérique et algorithme

Devant les difficultés rencontrées pour explorer toutes les partitions possibles construites par assemblage des composantes connexes répertoriées, parfois trop nombreuses, nous avons mis au point un algorithme d'optimisation qui réduit l'espace des configurations possibles. D'un point de vue pratique, on considère un nombre réduit d'ensembles de niveau, entre $Q = 4$ et $Q = 8$ niveaux, obtenus par quantification uniforme ou non-uniforme (e.g. quantification entropique) de l'histogramme de l'image f . Quatre étapes de traitement sont ensuite enchaînées au sein de l'algorithme 1 présenté ci-après.

La complexité de l'algorithme dépend essentiellement du nombre Q d'ensembles de niveau permettant d'établir la liste des domaines (i.e composantes connexes). Afin de calculer une partition plus rapidement mais de manière sous-optimale, nous avons proposé un algorithme itératif qui revient à chaque itération à sélectionner le domaine dont la contribution abaisse l'énergie le plus favorablement. Ce domaine devient une région à préserver et est supprimé de la liste des domaines à traiter. En définitive, le nombre de partitions explorées est $\frac{P_{max}(P_{max}+1)}{2}$ au lieu de $2^{P_{max}-1}$. Cet algorithme présente l'avantage pratique de pouvoir être implanté de façon économe en temps de calcul et mène bien souvent au minimum global de l'énergie sous-jacente. Ceci dit, l'algorithme est une technique d'optimisation locale et ne peut pas atteindre systématiquement le minimum global (Kervrann & Trubuil 2002).

Les modèles d'énergie introduits par Beaulieu & Goldberg (1989) dans un cadre discret, ne permettent pas cependant de contrôler la régularité des frontières des régions, une propriété chère en traitement d'images et vision par ordinateur. Pour remédier à cela, il est toujours possible de convoluer « faiblement » l'image avec un noyau gaussien G_ρ et de segmenter la fonction $f_\rho = G_\rho \star f$. En contrôlant la régularité de l'image, on impose en effet une certaine longueur aux lignes de niveau. Dans un contexte très bruité, cette approche relativement simple nécessite d'adopter une démarche différente. Nous avons ainsi proposé d'introduire un filtrage non-linéaire de manière à conjointement lisser les lignes de niveau et réduire le bruit ambiant (Kervrann et al. 2002). L'algorithme complet de segmentation alterne les deux étapes suivantes : *i*) sélection non-itérative d'un sous-ensemble de lignes de niveau de l'image qui minimise l'énergie globale de segmentation ; *ii*) diffusion (Nordström 1990) sur les domaines progressivement identifiés et dont les frontières sont contraintes par la segmentation. Le lien entre l'extraction des lignes de niveau et le filtrage adaptatif est ainsi établi. À chaque étape, l'image est simplifiée en fonction des paramètres de diffusion adaptés à chaque domaine D_p . Il est alors possible de contrôler le nombre

Algorithme 1 (Segmentation d'images par sélection de lignes de niveau)

- *Initialisation* : Soit l'image originale f et un choix de paramètres : Q , λ et $|D_{min}|$.
 - *Segmentation* :
 1. Construire les Q ensembles de niveau.
 2. Étiqueter les composantes connexes des ensembles de niveau, établir la liste des domaines $\{D_1, \dots, D_{P_{max}}\}$ tel que $|D_p| \geq |D_{min}|$ et calculer $\{\bar{f}_1, \dots, \bar{f}_{P_{max}}\}$.
 3. Énumérer les $2^{P_{max}} - 1$ configurations indexées, une suite de bits exprimant la présence ou l'absence d'un domaine pour la configuration étudiée.
 4. Calculer les énergies associées aux différentes partitions et sélectionner la configuration optimale.
-

Algorithme 2 (Segmentation d'images par sélection de lignes de niveau et diffusion)

- *Initialisation* : Soit l'image $f = u_0$ et un choix de paramètres : Q , λ et $|D_{min}|$.
 - *Tant que* $u_k(x) \neq u_{k+1}(x), \forall x \in \Omega$ *répéter*
 - *Segmentation* :
 1. Construire les Q ensembles de niveau de u_k .
 2. Étiqueter les composantes connexes et établir la liste des domaines.
 3. Énumérer les configurations indexées, une suite de bits exprimant la présence ou l'absence d'un domaine pour la configuration étudiée.
 4. Calculer les énergies associées aux différentes partitions et sélectionner la configuration optimale.
 - *Lissage* :
 1. Actualiser les paramètres de diffusion sur chaque support.
 2. Réaliser une diffusion sur chaque support.
 - Incréments k
 - *Fin tant que.*
-

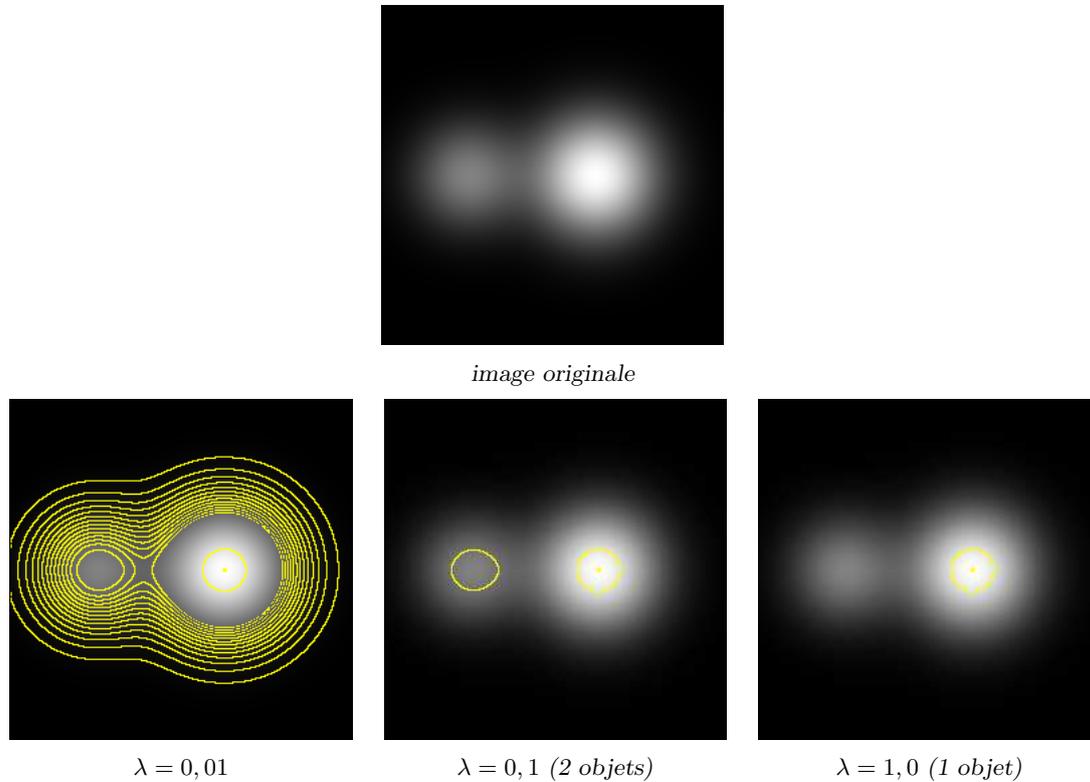


FIG. 2.3. Segmentation d'une image (256×256) artificielle correspondant à deux fonctions gaussiennes 2D de variance 792 et 1024 et dont les centres sont positionnés respectivement aux pixels (64 ; 128) et (160 ; 128). L'image est quantifiée uniformément avec $Q=32$ pas de quantification. Les segmentations sont obtenues en faisant varier λ (deuxième ligne).

d'itérations nécessaires pour obtenir la stabilité de la carte de segmentation. Invariablement, les mêmes lignes de niveau sont extraites si l'image restaurée u est suffisamment constante par morceaux, ce qui advient au bout de cinq ou dix étapes. L'algorithme 2 complet de segmentation est présenté ci-dessus. Les temps de calcul sont de l'ordre de quelques secondes en dimension 2.

Quelques résultats expérimentaux

La méthode décrite a été validée sur des images d'origine variée : imagerie satellitaire, imagerie médicale et microscopie 2D (Kervrann et al. 2002, Kervrann & Trubuil 2002, Kervrann 2001, Kervrann et al. 1999, Kervrann et al. 2000). Nous illustrons les performances de nos algorithmes avec quelques illustrations (Figs. 2.3, 2.4 et 2.5).

Au vu des résultats encourageants obtenus sur les images 2D, nous avons également conçu des schémas de segmentation d'images 3D pour la microscopie confocale (visualisation optique non-invasive de tissus en profondeur) tout aussi économes en temps de calcul et en mémoire. Cette

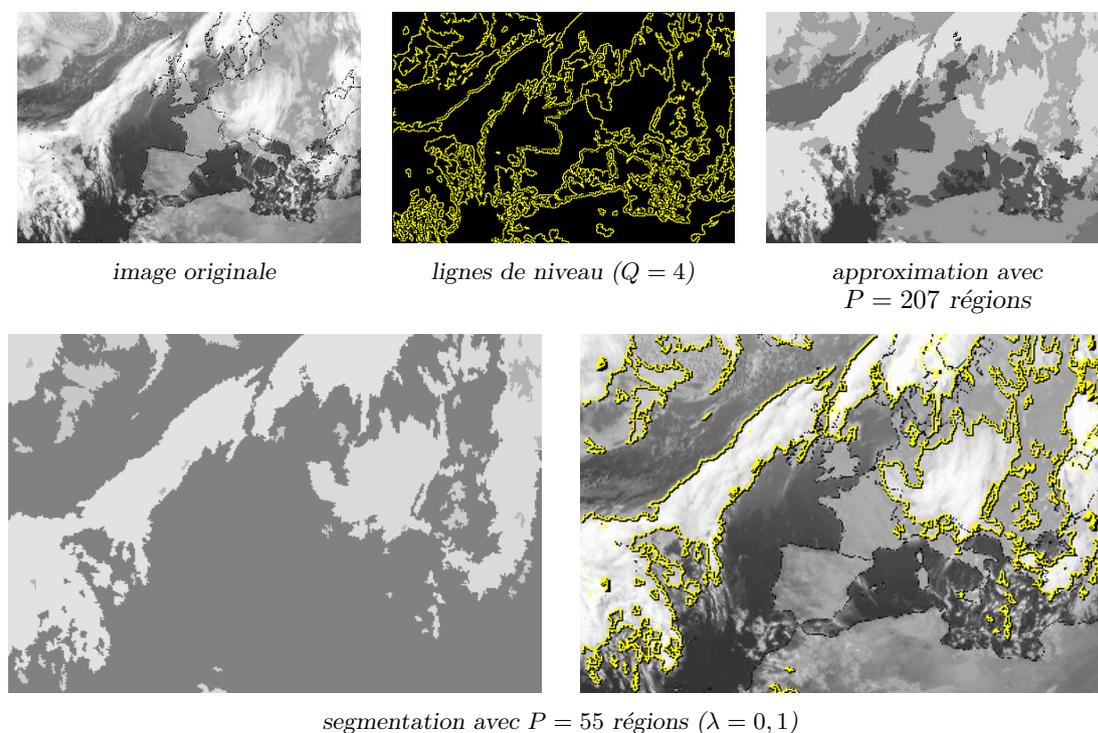


FIG. 2.4. Segmentation d'une image satellitaire (383×260). L'image est quantifiée uniformément avec $Q=4$ pas de quantification et les lignes de niveau associées aux 207 composantes connexes sont présentées sur la première ligne. La segmentation obtenue avec 55 composantes connexes est illustrée sur la seconde ligne ($\lambda = 0, 1$).

méthode de segmentation a été évaluée par plusieurs unités de l'Inra (Inra-ENVA « Biologie du Développement et Biotechnologies », Inra-UFDNH (Nantes), Inra-URPOI (Nantes), Inra-BioCell (Versailles)). Les algorithmes 2D et 3D correspondants ont été intégrés dans le logiciel Amira et dans le logiciel Quant-3D de l'unité Inra MIA Jouy-en-Josas.

2.3.3 Un autre éclairage : sélection de modèles, représentation parcimonieuse, modèles a contrario

Sélection de modèles

Dans un contexte uni-dimensionnel et gaussien, la détection de transitions abruptes dans un signal a fait l'objet de travaux récents en statistique. Contrairement aux approches passées, il s'agit d'évaluer la performance des estimateurs dans un cadre non-asymptotique, nécessitant une révision des termes de pénalisation généralement retenus pour contrôler l'approximation

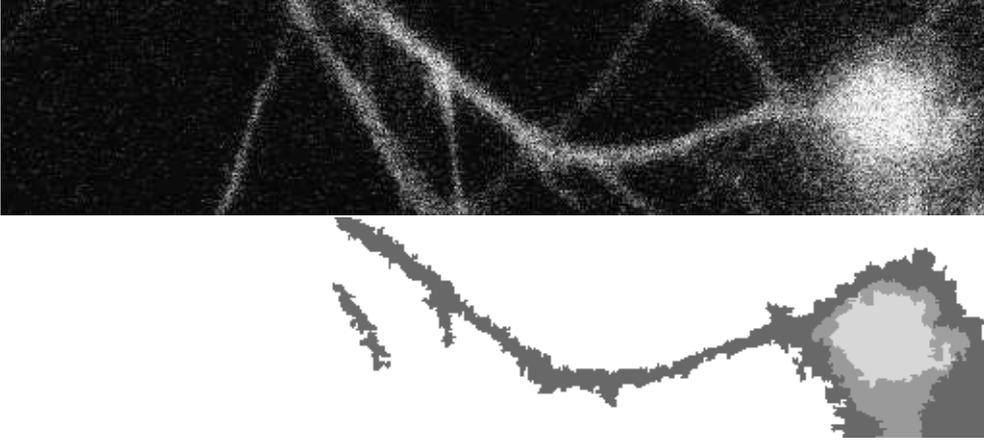


Fig. 2.5. Segmentation d'une cellule neuronale et diffusion en imagerie de microscopie 2D calcique (image 512×115) (Algorithme 2).

d'un signal. La modélisation du problème est donc très semblable à celle décrite dans (Birgé & Massart 2001, Lebarbier 2005, Huet 2006) et il s'agit toujours de déterminer des zones constantes sur un intervalle (ou un domaine en dimension 2) de longueur finie et égale à $N = |\Omega|$. Dans cette formulation du problème que nous adaptions au cas bidimensionnel, on considère le modèle statistique suivant :

$$f(x) = u(x) + \varepsilon(x), \quad x \in \Omega \quad (2.42)$$

où les erreurs $\{\varepsilon(x)\}$ sont des variables aléatoires i.i.d. gaussiennes de moyenne nulle et de variance σ^2 . La fonction $u : \Omega \rightarrow \mathbb{R}$ est supposée constante par morceaux.

En adoptant la présentation de Lebarbier (2005), à une segmentation (ou partition) du domaine Ω nous faisons correspondre un modèle $m \in \mathcal{M}_N$. Une segmentation m de dimension P_m s'écrit comme l'union des domaines $\{D_p\} : m = \cup_{p=1}^{P_m} D_p$. Soient $\mathcal{X} = \bigcup_{m \in \mathcal{M}_N} \mathcal{X}_m$ et \mathcal{X}_m un sous-espace linéaire de fonctions constantes par morceaux associé à un modèle m :

$$\mathcal{X}_m = \left\{ v = \sum_{p=1}^{P_m} v_p \mathbb{1}[x \in D_p] \right\} \quad (2.43)$$

où $v_p \in \mathbb{R}^{P_m}, p = \{1, \dots, P_m\}$. Une collection d'estimateurs $\{\hat{u}_m, m \in \mathcal{M}_N\}$ au sens des moindres-carrés est alors définie :

$$\begin{aligned} \hat{u}_m &= \arg \min_{v \in \mathcal{X}_m} \frac{1}{|\Omega|} \int_{\Omega} (f(x) - v(x))^2 dx \\ &= \sum_{p=1}^{P_m} \left(\frac{1}{|D_p|} \int_{D_p} f(x) dx \right) \mathbb{1}[x \in D_p]. \end{aligned} \quad (2.44)$$

Le problème d'estimation revient donc à déterminer le meilleur modèle \hat{m} et la dimension $\hat{P}_{\hat{m}}$ qui conduit à la « meilleure » segmentation du domaine Ω . L'estimateur \hat{u}_m recherché établit un bon compromis entre la complexité de la représentation et l'approximation de f . Pour atteindre cet objectif, nous introduisons le risque quadratique de l'estimateur considéré, qui se décompose usuellement en un terme de biais et un terme de variance :

$$\mathbb{E}[\|u - \hat{u}_m\|_N^2] = \|u - \bar{u}_m\|_N^2 + \frac{P_m}{N} \sigma^2 \quad (2.45)$$

où \bar{u}_m est la projection orthogonale de u sur \mathcal{X}_m :

$$\bar{u}_m = \sum_{p=1}^{P_m} \left(\frac{1}{|D_p|} \int_{D_p} u(x) dx \right) \mathbb{1}[x \in D_p]. \quad (2.46)$$

Le résultat remarquable obtenu dans le cadre non-asymptotique par Birgé & Massart (2001) et repris dans (Lebarbier 2005, Huet 2006) conduit finalement à chercher la segmentation optimale \hat{m} qui minimise le critère de moindres-carrés pénalisé suivant :

$$\hat{m} = \arg \min_{m \in \mathcal{M}_N} \frac{1}{N} \int_{\Omega} (f(x) - \hat{u}_m(x))^2 dx + \underbrace{\frac{P_m}{N} \sigma^2 \left(c_1 \log \left(\frac{N}{P_m} \right) + c_2 \right)}_{\text{pen}_N(m)}. \quad (2.47)$$

Le risque de l'estimateur obtenu après minimisation de ce critère est borné inférieurement et s'approche au mieux du risque idéal : $\mathbb{E}[\|u - \hat{u}_{\hat{m}}\|_N^2] \leq C_1 \inf_{m \in \mathcal{M}_N} \mathbb{E}[\|u - \hat{u}_m\|_N^2]$, $C_1 > 0$.

Le problème de segmentation d'images initialement posé dans la Section 2.3.2, est alors équivalent à un problème de « sélection de modèles » :

$$\hat{m} = \arg \min_{m \in \mathcal{M}_N} \frac{1}{|\Omega|} \sum_{p=1}^P \int_{D_p} \left(f(x) - \frac{1}{|D_p|} \int_{D_i} f(x) dx \right)^2 dx + \lambda \frac{P_m}{|\Omega|}. \quad (2.48)$$

En posant $\lambda = 2\sigma^2$ et $\lambda = \log(|\Omega|)\sigma^2$, l'énergie de segmentation que nous avons considérée coïncide avec le critère BIC (« Bayesian Information Criterion ») (Schwarz 1978a, Akaike 1978) et le critère C_p Mallows (Mallows 1974) étudiés depuis longtemps en « sélection de modèles » dans le cadre asymptotique. Une perspective assez simple sera donc d'adopter des pénalisations établies par Birgé & Massart (2001), dans le cadre non-asymptotique.

Représentation parcimonieuse

L'énergie (2.30) que nous avons considérée pour la segmentation d'images, trouve un intérêt nouveau pour peu qu'on s'intéresse au cas simple $\text{pen}(P, D_1, \dots, D_P) = P$. En effet, cette pénalisation comptabilise le nombre de domaines nécessaires pour décrire l'image. Dans le cadre de la représentation parcimonieuse des images (Chen, Donoho & Saunders 2001, Elad & Aharon 2006a), le modèle (2.30) (voir aussi (2.47)) exprime l'idée de représenter de manière économe l'image à partir d'« atomes morphologiques » identifiés sur l'image elle-même. Avec une nouvelle écriture, le problème de la segmentation que nous avons étudié en choisissant en

particulier $\text{pen}(P, D_1, \dots, D_P) = P$, peut s'exprimer sous la forme

$$\hat{\alpha} = \arg \min_{\alpha} \|\mathbf{A}\alpha - f\|_2^2 + \lambda' \|\alpha\|_0 \quad \text{avec } \lambda' > 0, \quad (2.49)$$

ou de manière équivalente

$$\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_0 \quad \text{sous la contrainte } \|\mathbf{A}\alpha - f\|_2^2 \leq \epsilon, \quad (2.50)$$

où ϵ est l'erreur d'approximation maximum tolérée, $\alpha \in \{0, 1\}^{P_{max}}$, \mathbf{A} est une matrice de grande dimension ($\dim(\mathbf{A}) = |\Omega| \times P_{max}$) et $\|\alpha\|_0$ est une pseudo-norme qui comptabilise le nombre d'éléments non nuls du vecteur α . Dans la situation exposée dans la section 2.3.2, les éléments du dictionnaire $\mathbf{A} = (A_1, \dots, A_{P_{max}})$ sont définis à partir des domaines D_p préalablement identifiés à partir des ensembles de niveau de l'image f :

$$A_p(x) = \left(\frac{1}{|D_p|} \int_{D_p} f(x) dx \right) \mathbf{1}[x \in D_p]. \quad (2.51)$$

Une perspective intéressante serait donc de considérer un intervalle continu $\alpha \in [0, 1]^{P_{max}}$ et de réviser la pseudo-norme $\|\alpha\|_0$ de manière à formuler le problème de segmentation comme un problème d'optimisation convexe.

Modélisation a contrario

Comme nous l'avons décrit dans ce chapitre, le problème de la détection de contours significatifs n'est pas étranger au problème de la segmentation d'images. Pour conforter cette idée, Desolneux, Moisan & Morel (2001) ont proposé de considérer toutes les lignes de niveau d'une image u , éventuellement obtenue après lissage de f par un noyau gaussien ($u = G_\rho \star f$), et d'associer à chaque ligne de niveau \mathcal{L} de longueur $|\mathcal{L}|$, un niveau de contraste minimum : $\mu = \min_{x \in \mathcal{L}} |Du(x)|$. L'approche *a contrario* évalue la probabilité que le contraste minimum mesuré sur une ligne de niveau \mathcal{L} soit supérieur à μ . Pour mener ce calcul, on considère la distribution empirique de l'image u afin d'exhiber les objets intéressants à détecter. Si on suppose que les variables $|Du(x)|$ sont indépendantes, le nombre de fausses alarmes associé à l'évènement que le contraste minimum mesuré sur une ligne de niveau de longueur $|\mathcal{L}|$ soit au moins égal à μ est relativement simple à calculer :

$$NFA(\mu, |\mathcal{L}|) = N_{\mathcal{L}} \cdot (H_u(\mu))^{|\mathcal{L}|} \quad \text{avec } H_u(\mu) = \frac{1}{|\Omega|} \sum_{x \in \Omega} \mathbf{1}[|Du(x)| \geq \mu], \quad (2.52)$$

où $N_{\mathcal{L}}$ désigne le nombre de lignes de niveau de u . Finalement, une ligne de niveau est dite « ϵ -significative » si $NFA(\mu, |\mathcal{L}|) \leq \epsilon$. En choisissant $\epsilon = 1$, on considère, qu'en moyenne, au plus une ligne de niveau est déclarée « significative » à tort parmi les lignes de niveau de l'image. Nous avons retenu ce principe dans le cadre d'une étude que nous avons menée en collaboration avec l'Ifremer et qui portait sur l'extraction de lignes significatives dans des images d'otolithes (Chessel, Fablet, Cao & Kervrann 2006, Chessel, Fablet, Kervrann & Cao 2008).

Cette description nous éclaire désormais un peu mieux sur les relations entre l'approche bayé-

sienne et variationnelle que nous avons étudiée et l'approche « fréquentiste » en segmentation d'images, si on accepte l'idée que la méthode *a contrario* évalue la fréquence d'apparition de lignes de niveau d'une certaine longueur dans l'image. On peut remarquer que ces deux approches visent à sélectionner une sous-collection de lignes de niveau de l'image issue de l'analyse de la distribution empirique de l'image.

2.4 Conclusion et perspectives

La mise en œuvre de méthodes bayésiennes et variationnelles peut s'avérer délicate en raison de la charge calculatoire nécessaire pour mener la minimisation globale des énergies non-convexes, même en dimension 2. Ces considérations nous ont conduit à proposer des fonctionnelles (terme d'attache aux données et régularisation) plus simples que les modèles habituellement préconisés, pour lesquelles les solutions minimisantes peuvent être confinées dans un domaine facilement explorable. Elles se déclinent par le choix de termes de régularisation contrôlant surtout l'aire des objets et le nombre d'objets à construire, adjoints à un terme d'attache aux données, ne reposant pas nécessairement sur un modèle de bruit blanc gaussien. Pour justifier l'emploi de ces termes d'énergie, on peut faire valoir que les courbes (ou surfaces) qui minimisent les fonctionnelles sous-jacentes se superposent à des lignes (ou surfaces) de niveau de l'image f , en nombre restreint et inconnu. Ces lignes (ou surfaces) de niveau, qui ont acquis un statut d'objet principal surtout en imagerie 2D et en morphologie mathématique, sont équivalentes aux frontières des composantes connexes des ensembles de niveau de l'image f , caractérisés par l'histogramme de f . Une image est ainsi décrite par une « mosaïque stochastique » de composantes connexes des ensembles de niveau de l'image, sans prise en compte de phénomènes d'occultation des objets. Cette description a permis d'unifier certaines notions de morphologie mathématique et de sélection automatique des seuils pour la segmentation. Il semble possible de poursuivre sur cette dynamique et de chercher à lisser des lignes de niveau selon d'autres mécanismes ou de s'intéresser aux ensembles de niveau d'une distribution de probabilité qui se déduirait d'attributs calculés sur l'image, par exemple en s'inspirant des travaux récents de Brox & Cremers (2009).

Les résultats que nous avons présentés sont originaux dans le sens où la plupart des résultats théoriques mettent en évidence l'existence d'un minimum global de la fonctionnelle étudiée, sans pour autant proposer toujours une méthode algorithmique rapide pour le déterminer. Ce fut moins le cas ces cinq dernières années où nous avons assisté à une quantité de méthodes de segmentation (markovienne ou variationnelle) pour lesquelles de nombreux algorithmes d'optimisation ont été proposés. Ce travail s'inscrit aussi dans un projet plus général de reconnaissance de forme, au sein duquel la question de savoir « quels critères optimiser » est devenue centrale. On peut considérer que la théorie des champs aléatoires conditionnels (CRF) ou champs aléatoires discriminants (DRF) (Kumar & Hebert 2006) permet aujourd'hui de mieux décrire les relations entre les modèles d'images et les observations. Une philosophie très différente de ce cadre de modélisation consiste justement à ne pas recourir à une description des composants de la scène. Le « principe d'Helmutz » (Desolneux, Moisan & Morel 2000, Desolneux et al. 2001) permet d'extraire des éléments significatifs permettant une compréhension de la scène. La segmentation de l'image par sélection de lignes de niveau significatives (au sens de la longueur et du contraste) est possible selon ce principe.

Approches non-paramétriques pour l'analyse d'images

3.1 Myopie, redondance et modèles d'images

Quelques approches en analyse d'images ont récemment été popularisées en raison de leur simplicité de mise en œuvre et compte tenu de l'amélioration spectaculaire des performances des ordinateurs. Il y a 10 ans, certains algorithmes auraient paru fantaisistes car peu économes en termes de calculs et en mémoire. Un exemple assez frappant est celui de la synthèse de textures proposée par Efros & Leung (1999), un algorithme qui a surclassé les approches paramétriques complexes issues d'une représentation par ondelettes ou inspirées des champs de Markov. La simplicité de cet algorithme est très surprenante et repose sur l'idée élémentaire de « copier » un motif de texture sans chercher à modéliser son contenu. Les modèles paramétriques statistiques capables de capter la variabilité observée sur un échantillon d'images sont bien moins efficaces que la procédure de « copier-coller » préconisée par Efros et Leung. Entre-temps, cet algorithme initial a reçu quelques améliorations (Criminisi, Pérez & Toyama 2004) et la stratégie pour synthétiser une image, apparemment empirique, a été justifiée dans un cadre plus formel d'estimation statistique (Levina & Bickel 2006). Un autre exemple très significatif est celui du débruitage d'images, un problème classique que nous avons évoqué dans le chapitre précédent. Dans un contexte très visité, Buades, Coll & Morel (2005) ont proposé de réduire le bruit dans une image en exploitant la redondance d'information captée par les motifs locaux qui la composent. L'algorithme de mise en œuvre des « moyennes non-locales » revient principalement à sélectionner dans toute l'image les motifs les plus similaires au motif de référence ; chaque pixel restauré coïncide ainsi avec la moyenne pondérée (par une distance entre motifs) des observations prises dans toute l'image. D'une manière générale, le bruit que l'on cherche à éliminer est supposé additif. En définitive, les deux algorithmes que nous venons d'évoquer ont pour caractéristique commune d'être moins « myopes » que leurs prédécesseurs. L'information n'est plus seulement analysée au niveau d'un pixel et de ses « quatre plus proches voisins » mais au-delà.

Le contexte spatial semble mieux capté par des motifs, et une modélisation paramétrique n'est plus indispensable. En définitive, si le débruitage ou la restauration d'images n'est pas encore un problème résolu, ceci est en partie dû à la nature des signaux-images qui contiennent de nombreux « objets » structurés et isolés, parfois apparentés à du bruit et donc difficiles à éliminer. Une voie plus récente en analyse d'images consiste à élaborer des approches non-paramétriques, guidées par les données, afin de décrire l'ensemble des régularités et géométries observées dans les images. Dans certains domaines scientifiques où l'imagerie est un outil d'investigation (e.g. biologie cellulaire), il est même recommandé de ne pas modéliser trop précisément l'image pour exhiber des informations peu attendues.

D'une manière générale, les méthodes d'analyse d'images les plus performantes sont généralement non-linéaires et peuvent s'appuyer sur la minimisation d'une fonctionnelle d'énergie, requérant un équilibrage des différents termes qui la composent. Ceci limite souvent leur usage et les paramètres sont finalement assez difficiles à comprendre pour un utilisateur occasionnel. Une méthode avec un ou deux paramètres serait acceptable si ceux-ci pouvaient être clairement ou au moins intuitivement compris. Sur le plan pratique, comment peut-on limiter le nombre de paramètres pour rendre l'algorithme acceptable? Exprimés sous la forme de « p-value », ces paramètres sont probablement à la portée de tous, mais peut-on l'exprimer ainsi à chaque fois? Dans un cadre bayésien, la gestion de ce problème est très présente à l'esprit. Une approche assez coûteuse consiste ainsi à « calibrer » les paramètres (Azencott 1987, Chalmond 2000), par exemple en fonction d'un critère et d'une base d'images d'apprentissage (Younes 2000), ce qui n'est pas toujours simple à réaliser. Pourtant, l'objectif essentiel est bien de proposer un algorithme efficace pour le problème posé, avec une faible intervention de l'opérateur si possible.

Enfin, caractériser la performance ou la fiabilité d'une solution produite par un algorithme d'analyse d'images reste encore un problème ouvert. Seules les méthodes de simulation de Monte-Carlo (Geman & Geman 1984) permettent réellement cela. Ces dernières années, nous avons abordé ce sujet mais pas de manière aussi frontale. Nous nous sommes intéressés à la performance des estimateurs exprimée par une mesure de risque ainsi qu'au niveau de confiance des solutions calculées. L'idée n'est plus seulement d'optimiser un critère afin de déterminer une solution telle que proposée dans le chapitre précédent, mais également de mesurer la fiabilité de cette solution en tout point de l'image, si possible.

Dans ce chapitre, nous abordons ces différentes problématiques et nous traitons les cas de la détection de changements temporels dans une paire d'images et du débruitage/restauration d'images.

3.2 Estimation non-paramétrique pour le débruitage d'images

Les méthodes de restauration d'images assez efficaces préservent les contours et prennent en compte de manière explicite la géométrie des objets (Starck et al. 2000, Tschumperlé 2006). Dans un cadre variationnel, les équations aux dérivées partielles sont désormais très populaires pour lisser les régions homogènes et préserver les zones de discontinuité spatiale de l'image (Perona & Malik 1990b, Rudin et al. 1992). Néanmoins, pour des raisons d'efficacité algorithmique, d'autres approches plus intuitives cherchent plus simplement à calculer une moyenne locale pondérée des valeurs d'intensité.

3.2.1 Filtrage de voisinage et méthodes non-linéaires

Assez faciles à mettre en œuvre et finalement assez efficaces, les filtres de voisinage (Lee 1983, Smith & Brady 1997, Tomasi & Manduchi 1998) ont connu un certain succès populaire et suscité un intérêt théorique. Soit une image bruitée $f : \mathcal{I} \subset \mathbb{R}^2 \rightarrow \mathbb{R}_+$ où $f(x_i), i = 1, \dots, N$, est l'intensité mesurée au pixel x_i et N représente le nombre de pixels de la grille discrète \mathcal{I} . Un exemple de filtre numérique de voisinage le filtre « bilatéral » (Tomasi & Manduchi 1998) :

$$\hat{u}(x_i) = \frac{\sum_{x_j \in \mathcal{I}} G_{h_r}(f(x_i) - f(x_j)) G_{h_s}(x_i - x_j) f(x_j)}{\sum_{x_j \in \mathcal{I}} G_{h_r}(f(x_i) - f(x_j)) G_{h_s}(x_i - x_j)} \quad (3.1)$$

où $\hat{u}(x_i)$ est la valeur restaurée au pixel $x_i \in \mathcal{I}$ à partir des observations et $G_{h_r}(\cdot)$ et $G_{h_s}(\cdot)$ sont des noyaux gaussiens de convolution d'écart-type h_r et h_s . La similarité entre les pixels x_i et x_j est évaluée par l'écart de niveau de gris $|f(x_i) - f(x_j)|$ et la distance spatiale entre les points x_i et x_j . Le filtre est principalement contrôlé par les paramètres h_r et h_s . En agissant sur h_r , le filtre est plus sensible aux contrastes locaux alors que h_s permet de contrôler la réduction de bruit et le flou dans l'image; ces deux paramètres nécessitent bien souvent d'être adaptés selon les images. Ces filtres gaussiens non-linéaires peuvent être reliés à des approches de filtrage également populaires en vision par ordinateur telles que la procédure « *mean-shift* », les M-estimateurs locaux, la diffusion non-linéaire et les approches régularisantes qui combinent à la fois des termes de lissage et des termes de fidélité aux données non-locaux (van de Weijer & van den Boomgaard 2001, Elad 2002, Barash 2002, Barash & Comaniciu 2004, Mrázek, Weickert & Bruh 2006, Sochen, Kimmel & Bruckstein 2001).

Aussi efficaces soient-elles, ces méthodes nécessitent toujours l'ajustement d'un nombre minimal de paramètres globaux qui contrôlent le lissage. Le problème est qu'il n'existe pas de solution très satisfaisante pour les déterminer à partir des données de l'image et, bien souvent, seule l'impression visuelle aide en pratique à les fixer (Tomasi & Manduchi 1998, Barash 2002). Par ailleurs, lorsque les caractéristiques locales varient spatialement, estimer un paramètre global de lissage est probablement mal venu. Dans cette situation, il est plus pertinent de déterminer l'échelle locale des caractéristiques de l'image, par exemple via une représentation espace-échelle (Lindeberg 1998). Parmi les travaux les plus récents, on peut citer ceux de Brox & Weickert (2004) qui ont suggéré d'analyser le flux de la « Variation Totale » afin de déterminer l'échelle locale et de discriminer les textures. Plus simplement, la variance locale est parfois suffisante pour adapter spatialement la régularisation (Black et al. 1998a, Gilboa, Sochen & Zeevi 2006). Enfin, nous pouvons mentionner les approches qui visent à estimer simultanément la composante stochastique (bruit), la composante oscillante (texture) et la composante à variation bornée de l'image (Osher, Sole & Vese 2002, Aujol et al. 2005). Difficiles à mettre en œuvre, ces méthodes ont néanmoins un potentiel indéniable en restauration d'images.

Une approche assez populaire ces dernières années consiste à exploiter l'information contextuelle locale résumée dans des motifs locaux pour capturer toutes les « formes » observées dans les images pour des applications en synthèse (Efros & Leung 1999), en reconstruction (Freeman et al. 2000, Fitzgibbon, Wexler & Zisserman 2003, Jovic, Frey & Kannan 2003, Criminisi et al. 2004) et en restauration d'images (Roth & Black 2005, Awate & Whitacker 2006). Bien souvent le cadre de modélisation choisi est le cadre bayésien, et une phase d'apprentissage

est nécessaire pour résumer toute l'information (Freeman et al. 2000, Roth & Black 2005). Le filtre des « moyennes non-locales » (Buades et al. 2005) reprend aussi cette idée pour le débruitage d'images. L'intensité restaurée au pixel $x_i \in \mathcal{I}$ est la moyenne pondérée des niveaux de gris mesurés dans toute l'image :

$$NL f(x_i) = \frac{1}{C(x_i)} \sum_{x_j \in \mathcal{I}} w(x_i, x_j) f(x_j) \quad (3.2)$$

où $C(x_i)$ est une constante de normalisation, i.e. $C(x_i) = \sum_{x_j \in \mathcal{I}} w(x_i, x_j)$. Les poids $w(x_i, x_j)$ sont définis comme :

$$w(x_i, x_j) = \exp \left(-\frac{1}{h^2} \sum_{\nu \in \mathbb{R}^2} G_a(s) |f(x_i + \nu) - f(x_j + \nu)|^2 \right) \quad (3.3)$$

et expriment le degré de similarité entre les motifs de la paire de pixels x_i and x_j . Le paramètre h permet de contrôler ce filtre qui peut être considéré comme la version étendue du filtre (3.1) qui prend en compte la distance spatiale entre les pixels de l'image. Un noyau de convolution $G_a(\cdot)$ d'écart-type a permet de contrôler la taille des motifs impliqués dans le calcul. Nous avons proposé d'améliorer les performances de ce filtre, en estimant de manière locale et adaptative le voisinage optimal autour de chaque pixel de l'image (Kervrann & Boulanger 2006a, Kervrann & Boulanger 2008).

3.2.2 Estimation adaptative et équilibre biais-variance

L'idée directrice est de sélectionner, dans un voisinage spatial adapté pour chaque pixel, des motifs qui sont des copies légèrement modifiées du motif de référence, centré au pixel considéré. Pour estimer ce voisinage, la stratégie retenue consiste à faire croître une fenêtre autour de chaque point de l'image, tant qu'un compromis entre le biais et la variance de l'estimateur à construire n'est pas obtenu. Cette adaptation est similaire à la procédure d'estimation non-paramétrique proposée à l'origine par Lepski (1991). L'algorithme mis en œuvre consiste à calculer des moyennes locales dans un voisinage adapté en s'appuyant sur des motifs locaux pour pondérer les observations. Il se compare très favorablement aux filtres gaussiens non-linéaires, aux techniques ondelettes les plus sophistiquées (Donoho & Johnstone 1994, Starck et al. 2000, Portilla, Strela, Wainw. & Simoncelli 2003, Pizurica & Philips 2006, Le Pennec & Mallat 2005, Dabov, Foi & Egiazarian 2007) voire aux approches parcimonieuses très récentes (Elad & Aharon 2006a, Mairal et al. 2008, Mairal, Bach, Ponce, Sapiro & A. 2009). L'algorithme, dirigé par les données, est donc d'une grande simplicité et nécessite l'ajustement d'un ou deux paramètres seulement (Kervrann & Boulanger 2006a, Kervrann & Boulanger 2008).

Formulation du problème

En chaque point x_i de l'image, nous observons une variable bruitée $f(x_i) \in \mathbb{R}_+$. Les observations sont réparties régulièrement sur une grille \mathcal{I} discrète et bidimensionnelle :

$$f(x_i) = u(x_i) + \varepsilon(x_i), \quad i = 1, \dots, N \quad (3.4)$$

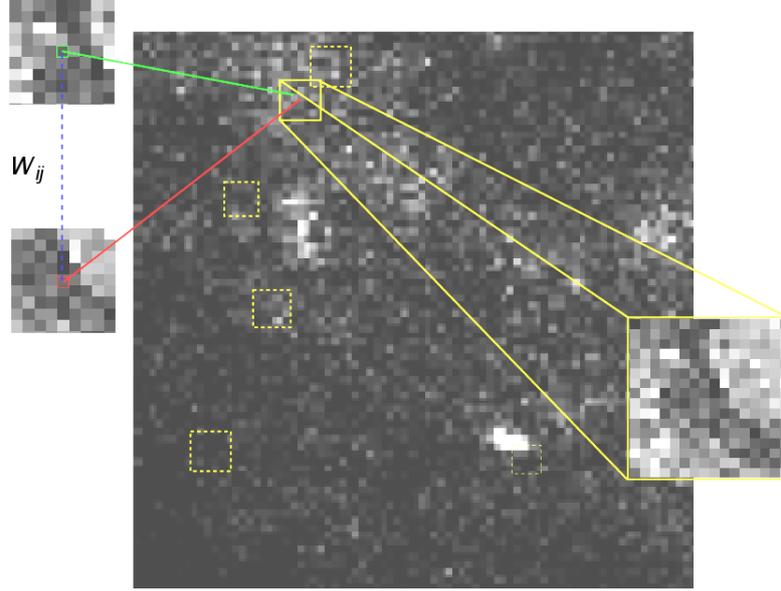


FIG. 3.1. Illustration de la sélection de motifs dans un voisinage local sur une image de microscopie optique (fluorescence).

où $x_i \in \mathcal{I} \subset \mathbb{R}^2$ représente les coordonnées spatiales de l'image et N est le nombre de pixels de l'image. Nous supposons que les erreurs $\varepsilon(x_i)$ sont des variables aléatoires gaussiennes, centrées, indépendantes et de variance σ^2 .

Afin de reconstruire $u : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ à partir des observations, il est nécessaire d'introduire des hypothèses minimales qui décrivent la structure de l'image. Nous supposons que les images reconstruites sont suffisamment régulières au sens où : $\sup_{x_i \in \mathcal{I}} u(x_i) < C_0$ et $\forall x_i, x_j \in \mathcal{I}, |u(x_i) - u(x_j)| \leq C_1$ (u est Lipschitz continue) avec $0 < C_0, C_1 < +\infty$. Nous postulons que la variable inconnue $u(x_i)$ peut être obtenue en calculant une moyenne pondérée des points d'observation situés dans un voisinage $V(x_i)$ local et variable spatialement, centré au point x_i . Les points $x_j \in V(x_i)$ qui repèrent des motifs locaux similaires au motif centré au point x_i participeront davantage que les autres points de l'image au calcul de cette moyenne locale. De manière implicite, cette description revient à faire l'hypothèse qu'il existe des répétitions du motif central dans le voisinage $V(x_i)$ du point x_i permettant de reconstruire la variable inconnue $u(x_i)$ (voir Fig.3.1).

Dans notre approche, un motif $\mathbf{u}(x_i)$ est simplement décrit par une fenêtre de taille $\sqrt{n} \times \sqrt{n}$ pixels, centrée au point x_i . Par la suite, $\mathbf{u}(x_i)$ désignera indifféremment un motif ou un vecteur composé de n éléments, les pixels du motif étant rangés selon un ordre lexicographique. La taille du motif est un paramètre libre, identique pour tous les pixels de l'image, et choisi de manière arbitraire. Typiquement des motifs 7×7 ou 9×9 permettent généralement de capturer l'ensemble des formes et des géométries observées dans la plupart des images naturelles.

Néanmoins, une question en suspens est la détermination de la taille et de la forme du voisinage $V(x_i)$, variable à chaque pixel, qui doit permettre de construire le meilleur estimateur possible.

Dans ce contexte, l'ensemble \mathcal{N}_V des voisinages candidats joue un rôle majeur. Pour des raisons de simplicité et de complexité calculatoire, cet ensemble est restreint ici à une collection de fenêtres emboîtées

$$\mathcal{N}_V = \{\widehat{V}_k(x_i) : |\widehat{V}_k(x_i)| = (2^k + 1) \times (2^k + 1), k = 1, \dots, K\},$$

où $|\widehat{V}_k(x_i)| = \#\{x_j \in \widehat{V}_k(x_i)\}$ désigne le nombre de points dans $\widehat{V}_k(x_i)$ et K le cardinal de l'ensemble \mathcal{N}_V . Cette collection respectera également les conditions suivantes : $\widehat{V}_k(x_i)$ est centrée au point x_i et $\widehat{V}_k(x_i) \subset \widehat{V}_{k+1}(x_i)$. La fenêtre optimale, sélectionnée dans l'ensemble \mathcal{N}_V , minimise alors le « risque local quadratique » de l'estimateur de $u(x_i)$. Cette optimisation sera menée en faisant croître $\widehat{V}_k(x_i)$ à chaque itération k jusqu'à l'établissement d'un compromis entre le biais et la variance de cet estimateur qui correspond au minimum du critère.

Une procédure itérative pour une meilleure adaptation

La procédure que nous avons proposée est itérative et opère de la manière suivante (Polzehl & Spokoiny 2000) (Kervrann 2004, Kervrann & Boulanger 2006a, Kervrann & Boulanger 2008). À l'initialisation, nous choisissons une fenêtre locale $V_0(x_i)$ contenant seulement le point d'estimation x_i ($|V_0(x_i)| = 1$). Un premier estimateur $\widehat{u}_0(x_i)$ (et sa variance associée $\widehat{\tau}_0^2(x_i) = \text{Var}[\widehat{u}_0(x_i)]$) est simplement obtenu : $\widehat{u}_0(x_i) = f(x_i)$ et $\widehat{\tau}_0^2(x_i) = \widehat{\sigma}^2$, où un estimateur $\widehat{\sigma}^2$ de la variance du bruit remplace σ^2 puisque celle-ci n'est pas nécessairement connue. À l'itération suivante, on considère une fenêtre $V_1(x_i)$ de taille supérieure, centrée au point x_i et telle que $V_0(x_i) \subset V_1(x_i)$. Chaque point $x_j \in V_1(x_i)$ reçoit un poids $w_1(x_i, x_j)$ qui mesure la similarité entre deux motifs $\widehat{\mathbf{u}}_0(x_i)$ et $\widehat{\mathbf{u}}_0(x_j)$, obtenus à la première itération. On rappelle que la taille n du motif est fixe pour tous les points de l'image. Les points x_j pour lesquels les motifs progressivement restaurés $\widehat{\mathbf{u}}_0(x_j)$ sont des copies altérées du motif central $\widehat{\mathbf{u}}_0(x_i)$ reçoivent un poids significatif proche de 1, et 0 dans le cas contraire. Nous définissons alors l'estimateur $\widehat{u}_1(x_i)$ comme la moyenne pondérée des points d'observation $f(x_j)$ situés dans le voisinage $V_1(x_i)$. Nous continuons ainsi en faisant croître à chaque itération k la fenêtre $V_k(x_i)$ autour de chaque point x_i tant que $k \leq K$ où K désigne le nombre maximal d'itérations de l'algorithme. À chaque itération $k \geq 1$, l'estimateur $\widehat{u}_k(x_i)$ au sens du maximum de vraisemblance (MV) et sa variance $\widehat{\tau}_k^2(x_i)$ sont donc de la forme

$$\widehat{u}_k(x_i) = \sum_{x_j \in V_k(x_i)} w_k(x_i, x_j) f(x_j), \quad \widehat{\tau}_k^2(x_i) = \widehat{\sigma}^2 \sum_{x_j \in V_k(x_i)} (w_k(x_i, x_j))^2 \quad (3.5)$$

où les poids $w_k(x_i, x_j)$ sont des variables continues et satisfont les conditions usuelles $0 \leq w_k(x_i, x_j) \leq 1$ et $\sum_{x_j \in V_k(x_i)} w_k(x_i, x_j) = 1$. Les poids sont obtenus en comparant les motifs restaurés $\widehat{\mathbf{u}}_{k-1}(x_i)$ et $\widehat{\mathbf{u}}_{k-1}(x_j)$, obtenus à l'itération $k - 1$:

$$w_k(x_i, x_j) \propto \exp\left(-\frac{\text{dist}(\widehat{\mathbf{u}}_{k-1}(x_i), \widehat{\mathbf{u}}_{k-1}(x_j))}{\lambda_\alpha}\right). \quad (3.6)$$

En raison de la décroissance du noyau, les distances importantes entre les motifs conduisent à des poids proches de 0, et λ_α agit comme un seuil de détection. Notons aussi que l'usage de poids permet de relâcher l'hypothèse structurelle qui décrit le voisinage comme une simple fenêtre

centrée au point considéré. La distance L_2 est suffisante comparer deux motifs locaux (Efros & Leung 1999, Criminisi et al. 2004, Buades et al. 2005). Néanmoins, nous avons privilégié la distance suivante

$$\text{dist}(\widehat{\mathbf{u}}_k(x_i), \widehat{\mathbf{u}}_k(x_j)) = (\widehat{\mathbf{u}}_k(x_i) - \widehat{\mathbf{u}}_k(x_j))^T \widehat{\Sigma}_k^{-1}(x_i, x_j) (\widehat{\mathbf{u}}_k(x_i) - \widehat{\mathbf{u}}_k(x_j)) \quad (3.7)$$

où $\widehat{\Sigma}_k^{-1}(x_i, x_j)$ est une matrice $n \times n$ diagonale qui prend en compte les variances $\widehat{\tau}_k^2(\cdot)$ des estimateurs obtenus pour chaque point du motif. L'hypothèse que $\widehat{\mathbf{u}}_{k-1}(x_i)$ soit une copie modifiée de $\widehat{\mathbf{u}}_{k-1}(x_j)$ est acceptée si cette distance est petite, i.e. $\text{dist}(\widehat{\mathbf{u}}_{k-1}(x_i), \widehat{\mathbf{u}}_{k-1}(x_j)) \leq \lambda_\alpha$. Le seuil $\lambda_\alpha > 0$ est alors choisi comme un quantile d'une distribution $\chi_{n,1-\alpha}^2$ avec n degrés de liberté, contrôlant la probabilité d'erreur de type I.

Avec ce choix adaptatif de la taille de la fenêtre et ce choix particulier de la fonction de poids qui prend en compte la similarité entre motifs locaux progressivement restaurés, l'estimateur proposé est non-linéaire. L'utilisation de fenêtres qui peuvent se recouvrir contribue par ailleurs à un très bon lissage de l'image, sans effet de bloc et sans artefact visuel indésirable. La question qui reste en suspens pour bon nombre de filtres de voisinage est la taille de la fenêtre qu'il s'agit également d'estimer en tout point de l'image.

Une fenêtre de lissage idéale

Nous abordons à présent le problème de la sélection optimale de la fenêtre $V_k(x_i)$ adaptée à chaque pixel x_i . Il est bien compris que, la régularité de l'image pouvant varier rapidement, une mesure globale évaluant le risque de notre estimateur est inappropriée pour refléter la précision ponctuelle. Nous avons donc adopté le « risque local quadratique » comme mesure de performance. Celui-ci se décompose usuellement en la somme du biais $\widehat{b}_k(x_i)$ au carré et de la variance $\widehat{\tau}_k^2(x_i)$:

$$[\mathbb{E}|\widehat{\mathbf{u}}_k(x_i) - u(x_i)|^2]^{1/2} = |\widehat{b}_k^2(x_i) + \widehat{\tau}_k^2(x_i)|^{1/2}. \quad (3.8)$$

Notre objectif est désormais d'optimiser ce critère vis-à-vis de la taille de la fenêtre de lissage $V_k(x_i)$, pour chaque pixel de l'image. En fait, la solution optimale explicite n'est pas connue pour un tel estimateur. On peut supposer qu'elle dépend néanmoins de la régularité de la fonction inconnue $u(x_i)$. Une idée naturelle est alors d'analyser le comportement du biais et de la variance lorsque $V_k(x_i)$ croît ou décroît avec k de la manière suivante :

- Le terme de biais $\widehat{b}_k(x_i) = \mathbb{E}[\widehat{\mathbf{u}}_k(x_i) - u(x_i)]$ est non aléatoire et caractérise la précision de l'approximation de la fonction u au point x_i . Comme il dépend explicitement de la fonction inconnue $u(x_i)$, son comportement est douteux. Néanmoins, si on utilise l'inégalité géométrique $|x_j - x_i| \leq \frac{\sqrt{2}}{2} |V_k(x_i)|^{1/2}$ et si on suppose qu'il existe une constante réelle $C_1 < \infty$ telle que $|u(x_j) - u(x_i)| \leq C_1 |x_j - x_i|$ (u est Lipschitz continue), alors $|\widehat{b}_k(x_i)| \leq \frac{C_1}{\sqrt{2}} |V_k(x_i)|^{1/2}$. Par conséquent, $|\widehat{b}_k(x_i)|^2$ est de l'ordre de $O(|V_k(x_i)|)$ et typiquement augmente lorsque $V_k(x_i)$ croît (voir (Katkovnik, Egiazarian & Astola 2002, Stankovic 2004)).
- Par définition, $0 \leq w_k(x_i, x_j) \leq 1$ et $\sum_{x_j \in V_k(x_i)} w_k(x_i, x_j) = 1$, ce qui nous permet d'écrire : $\frac{\widehat{\sigma}^2}{|V_k(x_i)|} \leq \widehat{\tau}_k^2(x_i) \leq \widehat{\sigma}^2$. Par ailleurs, on peut raisonnablement supposer qu'il existe une

constante $0 \leq \gamma^2 \leq 1$ telle que $\widehat{\tau}_k^2(x_i) \approx \widehat{\sigma}^2 |V_k(x_i)|^{-\gamma^2}$. Aussi, lorsque $V_k(x_i)$ augmente, davantage de points d'observation sont utilisés pour construire $\widehat{u}_k(x_i)$ et, de fait, $\widehat{\tau}_k^2(x_i)$ décroît.

En résumé, le biais et l'écart-type de l'estimateur local sont des fonctions monotones avec un comportement opposé. Afin d'approcher le minimum du « risque local quadratique », une démarche possible est de minimiser une borne supérieure de la forme suivante :

$$\mathbb{E}|\widehat{u}_k(x_i) - u(x_i)|^2 \leq \frac{C_1^2}{2} |V_k(x_i)| + \frac{\widehat{\sigma}^2}{|V_k(x_i)|^{\gamma^2}}.$$

La solution explicite $|V^*(x_i)| = \left[\frac{2\gamma^2 \widehat{\sigma}^2}{C_1^2} \right]^{\frac{1}{\gamma^2+1}}$ dépend des constante imprécises C_1 et γ mais nous permet de mettre en évidence que, dans le voisinage de la solution optimale $|V^*(x_i)|$, l'inégalité $|b^*(x_i)| \leq \gamma \tau^*(x_i)$ ne dépend pas de l'image (Katkovnik et al. 2002, Stankovic 2004). Par conséquent, une stratégie possible pour déterminer une fenêtre optimale est de faire croître la fenêtre $V_k(x_i)$ tant que $|b_k(x_i)|$ reste inférieur à $\gamma \widehat{\tau}_k(x_i)$, où $\gamma > 0$ est une constante à préciser. Cependant, une réelle difficulté est de mettre en place un tel critère de détection car le biais n'est pas observable et γ est inconnu. On peut adopter finalement la démarche suivante qui consiste à décomposer l'estimateur de la manière usuelle suivante $\widehat{u}_k(x_i) = u(x_i) + \widehat{b}_k(x_i) + e(x_i)$ où $e(x_i) \sim \mathcal{N}(0, \widehat{\tau}_k^2(x_i))$. Ceci nous permet d'écrire l'inégalité $|\widehat{u}_k(x_i) - u(x_i)| \leq \widehat{b}_k(x_i) + \varkappa \widehat{\tau}_k(x_i)$, qui est vérifiée avec une grande probabilité si $\varkappa > 0$ est choisi suffisamment grand. En définitive, nous avons

$$[\mathbb{E}|\widehat{u}_k(x_i) - u(x_i)|^2]^{1/2} \leq |\widehat{b}_k(x_i) + \varkappa \widehat{\tau}_k(x_i)| \leq |\widehat{b}_k(x_i)| + \varkappa \widehat{\tau}_k(x_i), \quad (3.9)$$

et l'index de la fenêtre « oracle » est définie comme

$$\sup\{k \in \{1, \dots, K\} : |\widehat{u}_k(x_i) - u(x_i)| \leq (\gamma + \varkappa) \widehat{\tau}_k(x_i)\}. \quad (3.10)$$

Le point crucial est que cette inégalité ne dépend plus du biais de l'estimateur mais seulement de la fonction inconnue $u(x_i)$. Dans le paragraphe suivant, nous avons proposé une règle pratique qui exploite cette définition de la fenêtre optimale.

Dans notre approche, nous disposons au point x_i d'une collection d'estimateurs unidimensionnels $\{\widehat{u}_0(x_i), \dots, \widehat{u}_K(x_i)\}$ ordonnée en direction des fenêtres de taille $|V_k(x_i)|$ croissante où $\widehat{u}_k(x_i)$ correspond au meilleur estimateur possible. Tous ces estimateurs sont en compétition et sont comparés deux à deux. Parmi tous les candidats $\widehat{u}_k(x_i)$, nous décidons en définitive de choisir celui qui possède la variance $\widehat{\tau}_k^2(x_i)$ la plus petite. Finalement, cela revient à choisir une fenêtre de lissage avec la règle locale suivante (Lepski 1991, Juditsky 1997, Lepski, Mammen & Spokoiny 1997, Goldenshluger & Nemirovski 1997) :

$$\widehat{k}(x_i) = \max\{k \in \{1, \dots, K\} : |\widehat{u}_k(x_i) - \widehat{u}_{k'}(x_i)| \leq \varrho \widehat{\tau}_{k'}(x_i), \text{ pour tout } 1 \leq k' < k\} \quad (3.11)$$

où $\varrho = (2\gamma + \varkappa)$. Cette règle de contrôle revient à choisir la fenêtre $V_k(x_i)$ la plus grande possible telle que les estimateurs $\widehat{u}_k(x_i)$ et $\widehat{u}_{k'}(x_i)$ ne sont pas trop différents, pour tout $1 \leq k' < k$. Par ailleurs, si un estimateur $\widehat{u}_{k'}(x_i)$ s'éloigne de l'ensemble des estimateurs précédents, on admet

Algorithme 3 (Débruitage d'images)

Initialisation : Soit l'ensemble $\{\alpha, p, N\}$ des paramètres initiaux. Calculer $\hat{\sigma}^2$, fixer ϱ et initialiser $\{\hat{u}_0(x_i), \hat{\tau}_0^2(x_i)\}$ pour chaque point $x_i \in \mathcal{I}$.

$k = 1$

Répéter

- Pour chaque point $x_i \in \mathcal{I}$

- pour chaque point $x_j \in V_k(x_i)$, calculer $w_k(x_i, x_j) = \exp\left(-\frac{\text{dist}(\hat{u}_{k-1}(x_i), \hat{u}_{k-1}(x_j))}{\lambda_\alpha}\right)$
- calculer

$$\hat{u}_k(x_i) = \sum_{x_j \in V_k(x_i)} w_k(x_i, x_j) f(x_j) \quad \text{et} \quad \hat{\tau}_k^2(x_i) = \hat{\sigma}^2 \sum_{x_j \in V_k(x_i)} (w_k(x_i, x_j))^2$$

- tester la fenêtre

$$\hat{k}(x_i) = \max\{k \in \{1, \dots, K\} : |\hat{u}_k(x_i) - \hat{u}_{k'}(x_i)| \leq \varrho \hat{\tau}_{k'}(x_i), \text{ pour tout } 1 \leq k' < k\}.$$

Si cette règle est violée à l'itération k , refuser $\hat{u}_k(x_i)$ et prendre $\hat{u}_{k-1}(x_i)$ comme estimateur final au point x_i : $\hat{u}(x_i) = \hat{u}_{k-1}(x_i)$ et $\hat{k}(x_i) = k - 1$.

- incrémenter k

tant que $k \leq K$.

que le biais est déjà trop grand et que la taille de la fenêtre $V_k(x_i)$ est surestimée. Cette idée de construction davantage décrite dans (Kervrann & Boulanger 2008), a été proposé à l'origine par Lepski (1991).

L'ingrédient principal de la procédure que nous venons de décrire repose sur la construction d'une séquence de voisinages emboîtés. À l'initialisation, nous fixons $|V_0(x_i)| = 1$ ainsi que la taille du motif $\sqrt{n} \times \sqrt{n}$. Comme nous l'avons précédemment évoqué, la procédure de restauration s'appuie sur l'estimation préliminaire de la variance du bruit $\hat{\sigma}^2$ estimée ici de manière robuste à partir des données (Gasser, Sroka & Jennen-Steinmetz 1986, Polzehl & Spokoiny 2000, Kervrann & Trubuil 2004). L'algorithme final présenté ci-après est donc d'une grande simplicité et d'une grande efficacité comme en ont attesté les expérimentations menées et décrites dans (Kervrann & Boulanger 2006b, Kervrann & Boulanger 2006a, Kervrann & Boulanger 2008). Sa complexité est de l'ordre de $n \times |\mathcal{I}| \times (|V_1(\cdot)| + \dots + |V_N(\cdot)|)$ si l'image contient $|\mathcal{I}|$ pixels.

Propriétés de l'estimateur pour contrôler les paramètres

Dans la procédure précédente, le paramètre ϱ joue un rôle essentiel qu'il s'agit de contrôler pour une bonne adaptation. Afin de répondre à cette exigence, nous avons prouvé les propositions suivantes (Kervrann & Boulanger 2008) :

Proposition 1 L'évènement $\{\widehat{k}(x_i) = k\}$ au pixel x_i est vérifié avec une probabilité

$$\mathbb{P}(\widehat{k}(x_i) = k) \leq \sum_{k' < k} 2 \exp\left(-\frac{\varrho^2}{2}\right).$$

Il suit que ϱ doit être choisi tel que :

$$\varrho \leq \sqrt{2 \log \frac{K(K-1)}{1 - \mathbb{P}(\widehat{k}(x_i) \geq K)}}. \quad (3.12)$$

Sur le plan pratique, $\mathbb{P}(\widehat{k}(x_i) \geq K)$ est obtenu empiriquement à partir des contrastes locaux mesurés sur l'image originale et la variance du bruit.

Avec une telle procédure d'adaptation, nous avons mis en évidence que le risque quadratique local de l'estimateur est borné de la manière suivante :

Proposition 2 Si et seulement si $k^*(x_i) \leq \widehat{k}(x_i)$, il existe un estimateur adaptatif $\widehat{u}(x_i)$ dont la précision est celle du risque idéal à un facteur multiplicatif $\left[\frac{(2\gamma+\varkappa)}{\sqrt{1+\gamma^2}} + 1\right]^2$ près.

Par ailleurs, nous ne sous-estimons presque jamais la taille de la fenêtre d'estimation. Pour cela, il suffit de choisir ϱ dans l'intervalle $[2, 3]$ et un nombre d'itérations relativement faible (e.g. $K = 4$). Il suit que :

Proposition 3 L'évènement $\{\widehat{k}(x_i) < k^*(x_i)\}$ se produit avec une probabilité

$$\mathbb{P}(\widehat{k}(x_i) < k^*(x_i)) \leq k^*(x_i)(k^*(x_i) - 1) \exp\left(-\frac{\varrho^2}{2}\right).$$

L'ensemble de ces propositions permet en définitive de contrôler la qualité et la performance de l'estimateur et d'éviter à un utilisateur de régler des paramètres déterminés par l'image elle-même.

3.2.3 Un autre éclairage statistique

L'estimateur adaptatif et itératif que nous avons proposé pour réduire le bruit dans les images peut également se justifier dans un cadre variationnel (Boulanger, Kervrann, Bouthemy, Elbau, Sibarita & Salamero 2010) puis bayésien (Kervrann, Boulanger & Coupé 2007). Nous décrivons ici les contributions qu'il nous semble important de mentionner.

Une formulation dans un cadre variationnel

Des travaux relativement récents ont porté sur l'analyse de fonctionnelles non-locales permettant de restaurer les images dans un cadre discret (Elmoataz, Lezoray & Bougleux 2008) et dans un cadre continu (Gilboa & Osher 2007) que nous évoquons désormais. Soit une image continue

3.2 Estimation non-paramétrique pour le débruitage d'images

$u : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ et une image observée $f : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$. La fonctionnelle de Kindermann, Osher & Jones (2005) peut-être considérée comme le point de départ de nos investigations :

$$E(u, f) = \frac{\lambda}{\sigma^2} \|u - f\|^2 + \int_{\Omega \times \Omega} g \left(\frac{|u(x) - u(y)|}{h_u} \right) \omega \left(\frac{\|x - y\|}{h_s} \right) dx dy. \quad (3.13)$$

Une formulation plus générale revient à introduire deux paramètres d'échelle $Q(\cdot)$ et $h(\cdot)$ qui dépendent de la position spatiale dans l'image et à l'étendre au cas des motifs dans l'image :

$$E(u, f) = \frac{\lambda}{\sigma^2} \|u - f\|^2 + \int_{\Omega \times \Omega} g \left(\int_{\Omega} G(t) \frac{(u(x+t) - u(y+t))^2}{Q(x+t, y+t)} dt \right) \omega \left(\frac{\|x - y\|}{h(x, y)} \right) dy dx \quad (3.14)$$

où $\omega : \mathbb{R} \rightarrow \mathbb{R}$ est un noyau de convolution, $g : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction dérivable (typiquement $g(z) = 1 - e^{-z}$) et $Q : \Omega \times \Omega \mapsto \mathbb{R}$ et $h : \Omega \times \Omega \mapsto \mathbb{R}$ sont des fonctions telles que $Q(x, y) = Q(y, x)$ et $h(x, y) = h(y, x)$. La minimisation de la fonctionnelle (3.14) conduit à la solution suivante :

$$\hat{u}(x) = \frac{f(x) + \frac{2\sigma^2}{\lambda} \int_{\Omega \times \Omega} A(x, y, z) u(y) dz dy}{1 + \frac{2\sigma^2}{\lambda} \int_{\Omega \times \Omega} A(x, y, z) dz dy} \quad (3.15)$$

avec

$$A(x, y, z) = \frac{G(z)}{Q(x, y)} \omega \left(\frac{\|x - y\|}{h(x - z, y - z)} \right) g' \left(\int_{\Omega} G(t) \frac{(u(x+t-z) - u(y+t-z))^2}{Q(x+t-z, y+t-z)} dt \right). \quad (3.16)$$

Une fonctionnelle d'énergie plus simple à minimiser est de la forme suivante

$$E(u, f) = \int_{\Omega} \left(\frac{\int_{\Omega} B(x, y, z) f(y) dz dy}{\int_{\Omega} B(x, y, z) dz dy} - u(x) \right)^2 dx \quad (3.17)$$

avec

$$B(x, y, z) = \frac{G(s)}{Q(x, y)} \omega \left(\frac{\|x - y\|}{h(x - z, y - z)} \right) g' \left(\mathbb{E} \left[\int_{\Omega} G(t) \frac{(f(x+t-z) - f(y+t-z))^2}{Q(x+t-z, y+t-z)} dt \right] \right). \quad (3.18)$$

Cette expression fait intervenir l'espérance de la distance entre motifs, ce qui améliore la reconnaissance dans un environnement bruité. Sous l'hypothèse usuelle d'un modèle de bruit blanc gaussien, i.e. $\mathbb{E}[f(x)] = u(x)$ et $\text{Var}(f(x)) = \sigma^2$, on déduit l'estimateur final :

$$\hat{u}(x) = \arg \min_u E(u, f) = \frac{\int_{\Omega \times \Omega} \tilde{B}(x, y, z) f(y) dz dy}{\int_{\Omega \times \Omega} \tilde{B}(x, y, z) dz dy} \quad (3.19)$$

avec

$$\tilde{B}(x, y, z) = \frac{G(z)}{Q(x, y)} \omega \left(\frac{\|x - y\|}{h(x - z, x - z)} \right) \times g' \left(\int_{\Omega} G(t) \frac{(u(x+t-z) - u(y+t-z))^2 + 2\sigma^2}{Q(x+t-z, y+t-z)} dt \right). \quad (3.20)$$

L'implémentation numérique de cette solution est équivalente à l'Algorithme 3 décrit précédemment, pour lequel le noyau $G(\cdot)$ est supposé rectangulaire (Boulanger et al. 2010). Nous avons

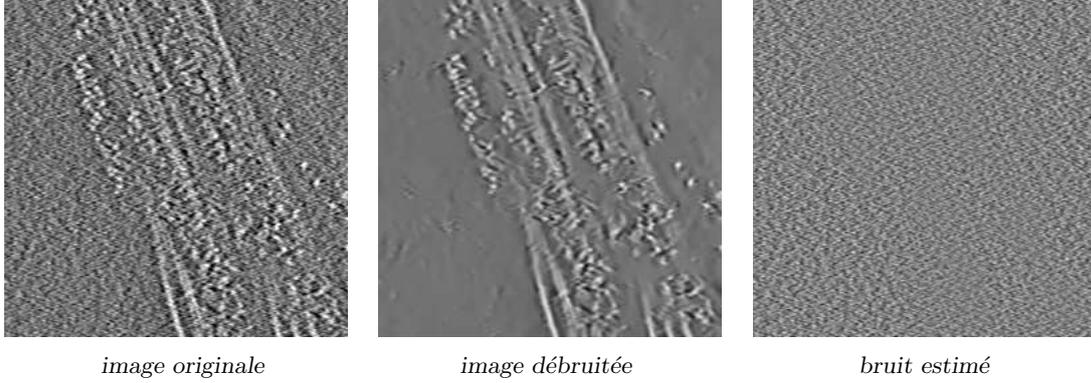


FIG. 3.2. Exemple de débruitage de microtubules et de protéines en cryo-tomographie électronique (Umr 6026 Cnrs Université Rennes 1).

également proposé d'estimer conjointement la taille de la fenêtre d'estimation, ce qui revient à estimer en tout point la fonction $h(x, y)$.

Le point de vue bayésien

Nous avons également cherché à établir les points communs pouvant exister entre le filtre des « moyennes non-locales » et la procédure que nous avons proposée. Pour cela, il est nécessaire d'adopter une modélisation qui s'inspire des filtres gaussiens non-linéaires établis dans ce contexte statistique.

Plus précisément, l'objectif est de restaurer directement un motif complet en adoptant une approche par blocs de pixels. Formellement, on cherche l'estimateur bayésien optimal $\hat{\mathbf{u}}(x_i) \in \mathbb{R}^n$, c'est-à-dire un motif de taille $\sqrt{n} \times \sqrt{n}$ centré au point $x_i \in \mathcal{I}, i = 1, \dots, N$, qui minimise un risque quadratique :

$$\begin{aligned} \mathbf{u}^*(x_i) &= \arg \min_{\hat{\mathbf{u}}(x_i)} \sum_{\mathbf{u}(x_i)} \|\mathbf{u}(x_i) - \hat{\mathbf{u}}(x_i)\|^2 p(\mathbf{u}(x_i) | \mathbf{f}(x_i)) \\ &= \frac{\sum_{\mathbf{u}(x_i)} \mathbf{u}(x_i) p(\mathbf{f}(x_i) | \mathbf{u}(x_i)) p(\mathbf{u}(x_i))}{\sum_{\mathbf{u}(x_i)} p(\mathbf{f}(x_i) | \mathbf{u}(x_i)) p(\mathbf{u}(x_i))} \end{aligned} \quad (3.21)$$

vis-à-vis de la distribution *a posteriori* $p(\mathbf{u}(x_i) | \mathbf{f}(x_i))$. Idéalement, il est nécessaire de spécifier la distribution *a priori* $p(\mathbf{u}(x_i))$ des motifs dans une image et la fonction de vraisemblance $p(\mathbf{f}(x_i) | \mathbf{u}(x_i))$. Une démarche assez naturelle serait d'introduire une méthode d'échantillonnage du type Monte-Carlo, mais elle est peu facile à mettre en œuvre dans notre contexte. En revanche, profitant de la redondance d'information observée aux pixels voisins, il est possible d'approcher l'estimateur de la manière suivante :

$$\hat{\mathbf{u}}_{N_V}(x_i) = \frac{\frac{1}{N_V(x_i)} \sum_{j=1}^{N_V(x_i)} \mathbf{u}(x_j) p(\mathbf{f}(x_i) | \mathbf{u}(x_j))}{\frac{1}{N_V(x_i)} \sum_{j=1}^{N_V(x_i)} p(\mathbf{f}(x_i) | \mathbf{u}(x_j))}, \quad (3.22)$$

où $N_V(x_i)$ désigne le nombre de motifs appartenant au voisinage local $V(x_i)$ du pixel x_i permettant d'établir un dictionnaire $\mathcal{D}(x_i) = \{\mathbf{u}(x_1), \mathbf{u}(x_2), \dots, \mathbf{u}(x_{N_V(x_i)})\}$. Implicitement, nous avons supposé une loi *a priori* uniforme sur ce dictionnaire en l'absence de connaissances plus précises. Nous reportons ainsi le problème sur la construction du dictionnaire et sur la modélisation de la distribution de probabilité $p(\mathbf{f}(x_i)|\mathbf{u}(x_j))$. Nous avons mis en évidence qu'il suffit de s'intéresser à la loi des distances entre motifs pour établir $p(\mathbf{f}(x_i)|\mathbf{u}(x_j))$. Dans un cas gaussien, $\text{dist}(\mathbf{f}(x_i), \mathbf{f}(x_j)) = \|\mathbf{f}(x_i) - \mathbf{f}(x_j)\|^2 / (2\sigma^2)$ suit une loi χ_n^2 avec n degrés de liberté. En considérant des motifs suffisamment grands de taille n , $\sqrt{2} \text{dist}(\mathbf{f}(x_i), \mathbf{f}(x_j))$ est approximativement distribuée suivant une loi normale de moyenne $\sqrt{2n-1}$ et de variance unitaire. Finalement, le filtre que nous avons établi ne dépend que de la taille des motifs considérés, un paramètre relativement intuitif à préciser par l'utilisateur :

$$ANL_{\sigma,n}\mathbf{f}(x_i) = \frac{\sum_{j=1}^{N_V(x_i)} \exp -\frac{1}{2} \left(\frac{\|\mathbf{f}(x_i) - \mathbf{f}(x_j)\|}{\sigma} - \sqrt{2n-1} \right)^2 \mathbf{f}(x_j)}{\sum_{j=1}^{N_V(x_i)} \exp -\frac{1}{2} \left(\frac{\|\mathbf{f}(x_i) - \mathbf{f}(x_j)\|}{\sigma} - \sqrt{2n-1} \right)^2}. \quad (3.23)$$

L'établissement des dictionnaires locaux est obtenu à partir de statistiques suffisantes (e.g. moyennes et variances locales) calculées préalablement sur l'image. Le filtre proposé revient à accepter ou refuser des motifs/échantillons sélectionnés dans un voisinage sur la base de statistiques « résumées ». L'ensemble des motifs sert ensuite à calculer une espérance *a posteriori*. Cette construction n'est pas sans rappeler une stratégie bayésienne récente (« *Approximate Bayesian Computation* » (Beaumont, Zhang & Balding 2002)) qui justement permet de simuler une loi *a posteriori* avec un algorithme du type « acceptation-rejet » lorsque la fonction de vraisemblance n'est pas connue précisément. Une illustration de débruitage avec ce filtre bayésien en cryo-tomographie électronique est présenté sur la figure 3.2.

L'intérêt du point de vue bayésien est de pouvoir aussi redéfinir un filtre adapté au modèle paramétrique de bruit si celui-ci est bien établi. Dans un contexte d'imagerie médicale et ultrasonore, nous avons ainsi adapté ce filtre à la réduction du bruit de chatoiement (« speckle ») pour lequel nous avons considéré un modèle d'images relativement simple (Kervrann et al. 2007) : $f_i = u(x_i) + \sqrt{u(x_i)} \varepsilon_i, \varepsilon_i \sim \mathcal{N}(0, \sigma^2)$. Dans ce cas précis, la variance du bruit dépend du niveau du signal. Le filtre adapté dans cette situation a été expérimenté avec succès sur des images d'échographie (Coupé, Hellier, Kervrann & Barillot 2008, Coupé et al. 2009).

Notons que d'autres interprétations probabilistes ont également été proposées par Louchet (2008), Azzabou (2008) et Deledalle, Denis & Tupin (2009). Nous mentionnons ici le fait que maximiser la loi jointe $p(\mathbf{u}(x_i))$ des variables qui composent le motif permet d'établir une relation avec le filtre des « moyennes non-locales » et l'algorithme UINTA (Awate & Whitacker 2006) qui examine l'entropie des motifs dans une image. On considère alors l'estimateur à noyaux de la densité de probabilité ainsi défini :

$$p(\mathbf{u}(x_i)) \approx \frac{1}{N_V(x_i)} \sum_{j=1}^{N_V(x_i)} G_h(\mathbf{u}(x_i), \mathbf{u}(x_j)) \quad (3.24)$$

où $G_h(\cdot)$ est un noyau gaussien de variance h . Pour restaurer le pixel x_i , on résout le problème d'optimisation suivant :

$$\hat{u}(x_i) = \arg \max_{u(x_i)} \sum_{j=1}^{N_V(x_i)} G_h(\mathbf{u}(x_i), \mathbf{u}(x_j)). \quad (3.25)$$

En définitive, la procédure itérative est analogue à un algorithme du type « *mean-shift* » (Fukunaga & Hostetler 1975) et la « solution à un pas » correspond au filtre des « moyennes non-locales » (Brox, Kleinschmidt & Cremers 2008).

3.2.4 Deux exemples d'application des méthodes adaptatives

Afin d'illustrer les performances de ces algorithmes de débruitage, les figures 3.3 et 3.4 présentent quelques résultats obtenus sur des images réelles pour lesquelles la dégradation ne correspond pas nécessairement à un bruit blanc gaussien. Une analyse quantitative des performances de ces algorithmes est présentée dans (Kervrann & Boulanger 2006a, Kervrann & Boulanger 2006b, Kervrann et al. 2007, Kervrann & Boulanger 2008).

Les approches non-paramétriques que nous venons de présenter ont connu un certain succès pour beaucoup de domaines d'application. J'en cite deux assez révélateurs.

Débruitage de séquences temporelles d'images 2D-3D

Le travail présenté dans ce paragraphe correspond aux travaux de thèse de J. Boulanger (2003-2007) que j'ai co-encadré (Boulanger, Kervrann & Boutheymy 2005b, Boulanger, Kervrann & Boutheymy 2005a, Boulanger et al. 2007, Boulanger, Sibarita, Kervrann & Boutheymy 2008, Boulanger et al. 2010). L'objectif était de proposer une méthode de filtrage spatio-temporel capable d'améliorer significativement le rapport signal-sur-bruit de séquences d'images ou de volumes acquises en microscopie à fluorescence. En effet, malgré les progrès en terme de résolution, le rapport signal-sur-bruit des images demeure souvent faible à cause de leur acquisition en régime de faible flux (comptage de photons). Par ailleurs, ces séquences présentent un grand nombre de petits objets assimilables à des particules/vésicules qui doivent être suivis d'une image à l'autre idéalement. Sans une étape de restauration, la vitesse de ces objets ne peut être calculée avec précision et fiabilité. Les trajectoires définissent des structures linéiques dans le volume spatio-temporel qu'il s'agit de préserver. Nous avons donc traité cette situation relativement complexe avec une méthode d'estimation adaptative conforme à l'Algorithme 3.

L'Algorithme 3 ne demande aucune connaissance *a priori* sur le signal et permet d'obtenir pour chaque pixel une mesure de la confiance. Par ailleurs, il ne repose pas sur la minimisation d'un critère régularisant et ne met pas en œuvre d'étape d'estimation du mouvement apparent. En effet, dans un contexte bruité et pour des objets petits voire rapides, pouvoir s'affranchir du calcul du flot optique bien souvent nécessaire dans les schémas d'estimation de séquences, représente un avancée importante. La restauration des séquences est ainsi réalisée grâce à une approche basée sur des fenêtres spatio-temporelles adaptatives selon la philosophie précédente. L'idée est toujours de sélectionner dans un voisinage adapté pour chaque pixel, des motifs qui sont des copies légèrement modifiées du motif centré au pixel considéré. La méthode de restauration cherche

3.2 Estimation non-paramétrique pour le débruitage d'images

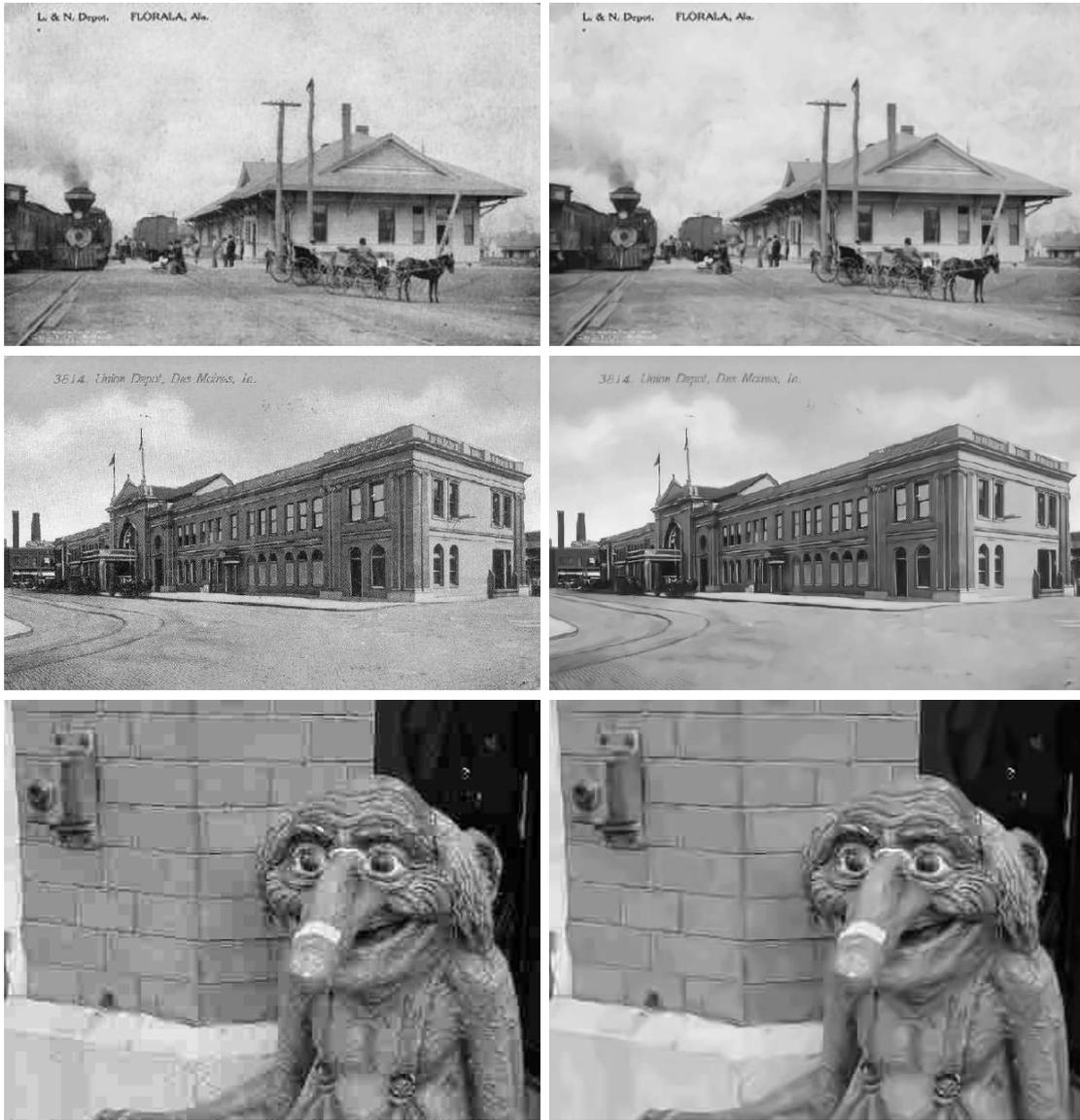


FIG. 3.3. Exemples de débruitage sur des images réelles dégradées par le vieillissement (première ligne), un effet de scanner (deuxième ligne) et une compression JPEG (troisième ligne).



FIG. 3.4. Exemple de restauration sur des images couleur réelles.

alors à calculer, en chaque point, une moyenne pondérée des observations sélectionnées dans un voisinage variable spatialement et temporellement. L'optimisation de la taille du voisinage repose sur un compromis biais/variance de l'estimateur, c'est à dire un bon équilibre entre précision de l'approximation et erreur stochastique. La méthode est itérative et s'adapte localement au degré de régularité de l'image. Elle permet ainsi de traiter des séquences temporelles de dimension élevée et ne requiert pas de dictionnaire, de technique d'apprentissage ni de recours à des transformées (e.g. bases d'ondelettes). Il suffit de découpler les variables spatiales et temporelle. La géométrie finale est donc plus complexe car la redondance est parfois plus significative dans le domaine spatial que dans le domaine temporel.

Les performances se sont révélées très satisfaisantes dans deux contextes bien différents. De manière assez surprenante, l'algorithme initialement conçu pour analyser des séquences de vidéo-microscopie, fut également très compétitif pour traiter des vidéos classiques pour lesquelles des

méthodologies ont été spécialement conçues (Boulanger et al. 2007). Entre-temps, des progrès très récents issus de ces résultats ont permis de franchir un palier important dans ce domaine. Dans le domaine de la vidéo-microscopie, nous avons réussi à diffuser et intégrer ces méthodes de débruitage sur des plate-formes expérimentales. L'intérêt principal réside dans l'espoir de capturer des événements dynamiques extrêmement rapides en limitant le temps d'exposition de l'échantillon. Ceci s'accompagne généralement d'une dégradation du signal qu'il s'agit de restaurer avec des méthodes numériques. Nous reviendrons sur ces aspects dans le chapitre 4.

Simplification d'images

Nous nous sommes également intéressés au problème de la simplification d'images et de la séparation des composantes texturales+bruit et des zones homogènes. Des travaux récents et relativement théoriques ont porté sur cette question (Osher et al. 2002, Aujol et al. 2005, Gilboa et al. 2006). L'approche que nous avons retenue est relativement simple et exploite le principe précédent (Kervrann 2004). L'idée est de considérer le modèle d'images suivant : $f(x_i) = u(x_i) + \varepsilon(x_i)$ où $x_i \in \mathcal{I}, i = 1, \dots, N$ et $\varepsilon(x_i) \sim \mathcal{N}(0, \sigma^2(x_i))$. On suppose que la variance du bruit blanc est désormais hétéroscédastique et dépend de la position spatiale dans l'image. Ceci permet de capter la variabilité relative à la fois au bruit et aux textures fines (relativement à la taille des fenêtres servant à l'estimation de la fonction u). Nous devons procéder dans un premier temps à l'estimation des variances locales. L'estimateur de Gasser et al. (1986) limité aux pixels situés dans un bloc de pixels, fut décisif dans ce cas :

$$\hat{\sigma}^2(x_i) = \frac{1}{|B(x_i)|} \sum_{x_j \in B(x_i)} r^2(x_i) \quad (3.26)$$

où $B(x_i)$ désigne un bloc de pixels centrés au point x_i et les pseudo-résidus se calculent très simplement à partir des observations $f(x_i) = f_{i_1, i_2} : r(x_i) = (2f_{i_1, i_2} - (f_{i_1+1, i_2} + f_{i_1, i_2+1}))/\sqrt{6}$.

Afin de détecter et éliminer les textures indésirables, il suffit d'introduire davantage de « myopie » dans la procédure en ne considérant plus les motifs pour sélectionner les points du voisinage mais uniquement les intensités mesurées ponctuellement en ces points. Le filtre obtenu est alors analogue au filtre « bilatéral » (3.1) pour lequel les paramètres dépendent de la position dans l'image. Expérimentalement, nous avons préconisé un noyau rectangulaire pour contrôler l'horizon spatial et un noyau choisi dans la famille des M-estimateurs (Huber) (Huber 1981) afin de détecter plus efficacement les contrastes locaux. En définitive, les ingrédients du filtre se limitent à ces simples expressions :

$$w_k(x_i, x_j) = \frac{\pi_k(x_i, x_j)}{\sum_{x_j \in V_k(x_i)} \pi_k(x_i, x_j)}, \quad \pi_k(x_i, x_j) = \begin{cases} 1 & \text{si } t_{k-1}(x_i, x_j) \geq 1 \\ t_{k-1}(x_i, x_j) & \text{sinon} \end{cases} \quad (3.27)$$

$$\hat{u}_k(x_i) = \sum_{x_j \in V_k(x_i)} w_k(x_i, x_j) f(x_j) \quad \text{et} \quad \hat{\tau}_k^2(x_i) = \sigma^2(x_i) \sum_{x_j \in V_k(x_i)} (w_k(x_i, x_j))^2 \quad (3.28)$$

où $0 \leq w_k(x_i, x_j) \leq 1$, $\sum_{x_j \in V_k(x_i)} w_k(x_i, x_j) = 1$ et $t_k(x_i, x_j) = \frac{\lambda \hat{\tau}_k(x_i)}{|\hat{u}_k(x_i) - \hat{u}_k(x_j)|}$. Cet algorithme itératif a été principalement décrit dans (Kervrann 2004) et un exemple de résultat est présenté

sur la figure 3.5.

Enfin, dans le cadre d'une collaboration avec une société spécialisée dans la conception de scanners de bio-puces, cet algorithme a été revisité afin de traiter des images de très grandes tailles « à la volée » (images 20000×10000 pixels). Les images de bio-puces considérées présentent des spots circulaires et sont relativement homogènes en intensité. Un traitement « par bandes » est possible avec cette méthode qui préserve très bien les discontinuités spatiales de l'image. La distribution de la taille des fenêtres et des poids, tout comme les estimations des variances des estimateurs locaux, furent des éléments essentiels pour analyser ces images.



image originale



image simplifiée



composante textuelle

FIG. 3.5. Simplification d'une image avec préservation des discontinuités spatiale et estimation la composante textuelle.

3.3 Détection non-paramétrique de changements dans une paire d'images

Dans la troisième partie de ce chapitre, nous abordons le thème de la comparaison d'images, un thème bien traité dans la littérature en analyse d'images et vision par ordinateur. Afin de mettre en correspondance deux vues d'une même scène à des fins d'interprétation ou de reconnaissance, il est bien souvent nécessaire d'identifier les pixels de l'image pour lesquels il n'est pas possible de trouver de correspondant dans une seconde image. Le problème de la détection de changements dans une paire d'images a pris encore plus d'importance ces dernières années notamment en raison des progrès technologiques incessants qui permettent d'observer de près (microscopie, image médicale) ou de loin (vidéo-surveillance, imagerie satellitaire (Kasetkasem & Varshney 2002)) des scènes relativement complexes en 2D ou 3D.

Un grand nombre de méthodes, supervisées et non-supervisées, ont été mises au point ces dix dernières années pour répondre aux exigences de certaines applications (Aach, Dumbgen, Mester & Toth 2001, Rosin 2002, Lisani & Morel 2003, Ince & Konrad 2005, Radke, Andra, Al Kofahi & Roysam 2005). La difficulté à surmonter consiste bien à distinguer les changements significatifs associés à des événements inhabituels observés dans la scène, de changements liés à des mouvements de caméra, des mouvements naturels dans la scène (e.g. mouvements d'arbres) ou du bruit inhérent au capteur. Les conditions d'éclairage de la scène peuvent également changer au cours du temps. Le problème devient donc particulièrement délicat à traiter si on souhaite définir une méthodologie robuste à tous ces facteurs environnementaux difficiles à modéliser. Les méthodes conventionnelles s'appuient d'une part sur une modélisation statistique de l'image de différence et sur la modélisation du contexte et des corrélations spatiales par champs de Markov (Kasetkasem & Varshney 2002, Benedek, Szirányi, Kato & Zerubia 2007). Dans le domaine de la vidéo qui suppose l'analyse de plusieurs dizaines d'images, un consensus a été obtenu. Profitant de la redondance temporelle, une approche simple et efficace consiste à modéliser la densité de probabilité de l'arrière-plan de la scène (supposé fixe) par un mélange de gaussiennes. La série d'images exploitée pour l'estimation des paramètres (algorithme EM) est dépourvue d'objets en mouvement (Elgammal, Harwood & Davis 2000, Jodoin, Mignotte & Konrad 2007). La détection de changements revient donc à tester chaque pixel de l'image connaissant un modèle paramétrique de la densité de probabilité de l'arrière plan (Stauffer & Grimson 1999). Dans le contexte plus délicat de l'estimation de champs de vecteurs de déplacement qui relie deux images, la détection des zones occultées pour lesquelles il n'y pas de correspondance possible, devient critique. Curieusement, peu de travaux ont porté sur ce point délicat (Heitz & Bouthemy 1993, Black & Fleet 2000, Ince & Konrad 2005) au regard des efforts considérables menés sur cette question dans un contexte assez voisin en stéréo (Kolmogorov & Zabih 2001, Sabater, Almansa & Morel 2008, Sun, Li, Kang & Shum 2005).

Compte tenu des différentes modélisations statistiques existantes, nous avons orienté nos efforts dans une direction quelque peu différente. En nous inspirant des travaux de Efros & Leung (1999) en synthèse d'images et ceux de Leung & Malik (1996) portant sur la détection d'éléments picturaux inhabituels dans une image, nous avons cherché à caractériser statistiquement les régions détectées. Les travaux que nous venons d'évoquer ont par ailleurs en commun le fait d'exploiter la propriété de redondance des images captée par des motifs locaux. Les travaux

récents en restauration d'images (Buades et al. 2005) reprennent également à leur compte cette propriété comme nous l'avons évoqué déjà dans ce chapitre. L'estimation de pixels occultés étant relativement ambiguë, nous associons à chaque motif de la première image plusieurs motifs observés dans la seconde image (Buades, Coll & Morel 2008) au lieu d'un seul comme cela est préconisé en estimation de champs de déplacement. Par la suite, nous collectons des décisions binaires prises dans le voisinage du pixel considéré, et ceci pour tous les pixels de l'image. En d'autres termes, il s'agit de comptabiliser le nombre de motifs voisins dans la seconde image, similaires au motif considéré dans la première image. Une décision finale est prise par agrégation des réponses binaires obtenues pour différentes tailles de motifs, selon un principe de fusion de décisions exposé par Niu, Varshney, Moore & Klamer (2004) et repris par Katenka, Levina & Michailidis (2008). On peut évoquer plusieurs avantages de ce mécanisme expérimenté dans un contexte de fusion de capteurs. Expérimentalement, il n'est nul besoin d'introduire de processus de régularisation pour obtenir une carte de détection régularisée ; le recouvrement des réponses dans un voisinage local couplé à une modélisation multi-échelle suffit pour garantir un résultat exploitable en analyse de scènes. Par ailleurs, on peut mentionner que le calcul de champs de déplacement est inutile et que la méthode est non-supervisée et ne repose sur aucune hypothèse forte sur la nature des images ou du bruit observé. Enfin, elle est relativement robuste aux légers mouvements de caméra assimilés à des vibrations ou à l'agitation dans l'arrière-plan de la scène. On peut donc s'abstenir, dans beaucoup de cas, de compenser le mouvement de la caméra (Ince & Konrad 2005). Les performances sur certaines paires d'images sont analogues à celles obtenues par un grand nombre de méthodes de détection de changements qui disposent d'une série complète d'images pour soustraire les objets mobiles.

La méthode que nous présentons maintenant (Kervrann, Boulanger, Pécot & Pérez 2009, Kervrann, Boulanger, Pécot, Pérez & J. Salamero 2009) plus en détail partage des points communs avec les méthodes basées sur des votes (Lisani & Morel 2003, Ince & Konrad 2005). D'une certaine façon, Aach et al. (2001) ont mis au point des tests statistiques (avec des hypothèses gaussiennes sous-jacentes) au niveau de chaque pixel et des fenêtres glissantes. Une approche plus intuitive est celle proposée par Shechtman & Irani (2007) qui caractérise l'auto-similarité captée par des motifs locaux dans les images. L'approche bayésienne préconisée par Boiman & Irani (2005) est également intéressante puisque l'idée est de détecter des situations dynamiques inhabituelles et observées dans une séquence entière d'images. Finalement nous reviendrons largement sur la méthode *a contrario* déjà évoquée dans le chapitre précédent qui a fait ses preuves également en détection de changements sur des paires d'images stéréo (Sabater et al. 2008). L'avantage certain est de pouvoir garantir qu'en moyenne, le nombre moyen de blocs mis en correspondance de manière erronée n'excède pas l'unité.

3.3.1 Le détecteur classique en traitement du signal

Soit une paire d'images $u_1 = (u_1(x_i))_{x_i \in \mathcal{I}}$ et $u_2 = (u_2(x_i))_{x_i \in \mathcal{I}}$, $i = 1, \dots, N$, définies sur une grille discrète rectangulaire $\mathcal{I} \subset \mathbb{R}^2$. Afin de détecter un changement au pixel x_i , on analyse deux blocs de pixels extraits dans chacune des deux images. On note $\mathcal{B}_1(x_i)$ et $\mathcal{B}_2(x_i)$ les deux blocs considérés de taille $\sqrt{n} \times \sqrt{n}$ et centrés au pixel x_i . Soit $\mathbf{d}_n(x_i) = \mathcal{B}_1(x_i) - \mathcal{B}_2(x_i) \in \mathbb{R}^n$ le vecteur de différence entre les deux blocs. Traditionnellement, on distingue les deux hypothèses

3.3 Détection non-paramétrique de changements dans une paire d'images

H_0 et H_1 suivantes :

$$\begin{aligned} H_0 : \mathbf{d}_n(x_i) &= \boldsymbol{\varepsilon}_n(x_i) \\ H_1 : \mathbf{d}_n(x_i) &= \mathbf{a}_n(x_i) + \boldsymbol{\varepsilon}_n(x_i) \end{aligned} \quad (3.29)$$

où $\mathbf{a}_n(x_i) \in \mathbb{R}^n$ est un vecteur connu et $\boldsymbol{\varepsilon}_n(x_i) \sim \mathcal{N}(0, \sigma^2 I_n)$ est vecteur gaussien de variance σ^2 connue.

Soient les probabilités de fausses alarmes et de détection définies respectivement de la manière suivante :

$$\begin{aligned} P_{FA}(x_i) &= \mathbb{P}(H_1 \text{ est choisie alors que } H_0 \text{ est vraie}), \\ P_D(x_i) &= \mathbb{P}(H_1 \text{ est choisie alors que } H_1 \text{ est vraie}). \end{aligned}$$

Le détecteur de Neyman-Pearson est communément utilisé pour distinguer les deux situations et vise à maximiser $P_D(x_i)$ sous la contrainte que $P_F(x_i) \leq \alpha$. Cette optimisation revient à effectuer un test d'hypothèses :

$$\Lambda(\mathbf{d}_n(x_i)) = \frac{f_1(\mathbf{d}_n(x_i))}{f_0(\mathbf{d}_n(x_i))} \underset{H_1}{\overset{H_0}{\leq}} \eta(x_i) \quad (3.30)$$

où $f_0(\mathbf{d}_n(x_i)) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{\|\mathbf{d}_n(x_i)\|_2^2}{2\sigma^2}}$ et $f_1(\mathbf{d}_n(x_i)) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{\|\mathbf{d}_n(x_i) - \mathbf{a}_n(x_i)\|_2^2}{2\sigma^2}}$. En définitive, on choisit η tel que :

$$P_{FA}(x_i) = \int_{\Lambda(\mathbf{d}_n(x_i)) > \eta(x_i)} f_0(\mathbf{d}_n(x_i)) d\mathbf{d}_n(x_i) = \alpha. \quad (3.31)$$

Par ailleurs, en prenant le logarithme du test d'hypothèses, on peut montrer que :

$$\begin{aligned} P_{FA}(x_i) &= \mathbb{P}(\langle \mathbf{d}_n(x_i), \mathbf{a}_n(x_i) \rangle > \gamma(x_i) | H_0) = Q\left(\frac{\gamma(x_i)}{\sigma \|\mathbf{a}_n(x_i)\|_2}\right) \\ P_D(x_i) &= \mathbb{P}(\langle \mathbf{d}_n(x_i), \mathbf{a}_n(x_i) \rangle < \gamma(x_i) | H_1) = Q\left(\frac{\gamma(x_i) - \|\mathbf{a}_n(x_i)\|_2^2}{\sigma \|\mathbf{a}_n(x_i)\|_2}\right) \end{aligned} \quad (3.32)$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire de deux vecteurs, $\gamma(x_i) = \sigma^2 \log(\eta) + \frac{1}{2} \|\mathbf{a}_n(x_i)\|_2^2$ et $Q(r) = \int_r^\infty e^{-r^2/2} dr$. En fixant $P_{FA}(x_i) = \alpha$, on obtient finalement :

$$\gamma(x_i) = \sigma \|\mathbf{a}_n(x_i)\|_2 Q^{-1}(\alpha) \quad \text{et} \quad P_D(x_i) = Q\left(Q^{-1}(\alpha) - \frac{\|\mathbf{a}_n(x_i)\|_2}{\sigma}\right). \quad (3.33)$$

Lors de la mise en œuvre, il est donc nécessaire de connaître la variance σ du bruit blanc gaussien et le vecteur $\mathbf{a}_n(x_i)$. Malheureusement, pour détecter des changements sur une paire d'images, on vérifie assez vite que l'hypothèse statistique de bruit blanc stationnaire sur toute l'image est peu réaliste. Par ailleurs, le vecteur $\mathbf{a}_n(x_i)$ est inconnu et dépend de la position spatiale dans l'image. Pour ces raisons, nous avons renoncé à cette modélisation et nous avons conçu un détecteur dont les performances sont néanmoins établies en termes de probabilités de fausses alarmes.

3.3.2 Principe de la fusion de décisions locales

En reprenant les notations précédentes, soit une paire d'images $u_1 = (u_1(x_i))_{x_i \in \mathcal{I}}$ et $u_2 = (u_2(x_i))_{x_i \in \mathcal{I}}$, $i = 1, \dots, N$, définies sur une grille discrète rectangulaire $\mathcal{I} \subset \mathbb{R}^2$. Afin de mesurer précisément la similarité entre ces deux images, l'idée que nous avons retenue revient à comptabiliser dans l'image u_2 le nombre de motifs similaires à un motif de référence de taille n sélectionné à la position x_i dans l'image u_1 . Le voisinage semi-local considéré, noté $B(x_i) \subseteq \mathcal{I}$, est de taille fixe et comporte $N_B = |B(x_i)|$ motifs observés aux positions x_j dans la seconde image. Pour simplifier la représentation, un motif est également un vecteur dont les éléments correspondent aux intensités lumineuses des pixels du motif (carré ou circulaire) rangées selon un ordre lexicographique. Théoriquement, l'espace de recherche peut être aussi large que voulu pour traiter le cas de la détection d'occlusions.

On propose de calculer une mesure de dissimilarité $\phi_{u_1 u_2}(x_i, x_j)$ non-négative définie comme

$$\phi_{u_1 u_2}(x_i, x_j) = \sum_{\nu \in \mathbb{R}^2} W_n(\nu) g(u_1(x_i + \nu), u_2(x_j + \nu)) \quad (3.34)$$

où $g : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}_+$ est une fonction mesurable choisie en fonction de l'application visée et $W_n(\cdot)$ permet d'accorder plus d'importance aux éléments du motif en fonction de leur distance spatiale au point central x_i . On choisit généralement $W_n(\nu) = \text{rect}_n(\nu) \triangleq \mathbf{1}[\|\nu\|_\infty \leq \frac{\sqrt{n}}{2}]$, ce qui revient à calculer la distance L_2 entre motifs de taille n . On considère que deux motifs ne sont pas similaires si le score excède un seuil $\varsigma(x_i)$ qui dépend de la position spatiale. Un choix assez intuitif pour identifier un changement significatif est de comptabiliser le nombre de décisions positives que l'on notera S_{N_B} , et de le comparer à un seuil global $T \in \{1, \dots, N_B\}$ où $N_B = |B(x_i)|$ désigne la taille de la fenêtre de sondage $B(x_i)$:

$$S_{N_B}(x_i) = \sum_{x_j \in B(x_i)} \mathbf{1}[\phi_{u_1 u_2}(x_i, x_j) \geq \varsigma(x_i)] \quad (3.35)$$

La décision $D(x_i) \in \{0, 1\}$ au pixel x_i est 1 si le nombre de décisions excède le seuil T (Niu et al. 2004, Katenka et al. 2008). Dans ce cas, nous déclarons qu'un changement est identifié au pixel x_i . La coopération entre pixels voisins permet de fiabiliser la prise de décision, surtout dans un contexte bruité. Dans cette procédure, les résultats sont très dépendants du choix du seuil T , de la taille n des motifs et de la taille N_B de la fenêtre de sondage. Afin de proposer un algorithme relativement simple à contrôler, nous avons proposé des critères pour choisir ces paramètres.

Dans notre scénario, la plupart des motifs se recouvrent puisque nous supposons $B(x_i)$ relativement petit. Par conséquent, $S_{N_B}(x_i)$ se définit comme la somme de variables aléatoires dépendantes de Bernouilli. Néanmoins, d'après le Théorème Central Limite pour des variables dépendantes, $S_{N_B}(x_i)$ est asymptotiquement gaussien. Sous l'hypothèse H_0 (pas de changement), l'espérance et la variance, à l'état stable, de $S_{N_B}(x_i)$ vérifient les conditions suivantes (Nemsick & Geraniotis 1992) : $\mathbb{E}[S_{N_B}(x_i)|H_0] = N_B p_{fa}(x_i, \varsigma(x_i))$ et $\text{Var}[S_{N_B}(x_i)|H_0] \leq N_B p_{fa}(x_i, \varsigma(x_i))(1 - p_{fa}(x_i, \varsigma(x_i)))$. D'après (Nemsick & Geraniotis 1992), il suit que $S_{N_B}(x_i)$ peut être approché par une somme de variables indépendantes et identiquement distribuées. Bien que les pixels voisins partagent de l'information, la probabilité $P_{fa}(x_i, T) \triangleq P(S_{N_B}(x_i) \geq T|H_0)$ qu'au moins T pixels

déclarent une décision positive sous H_0 , est donc donnée par la queue de la Binomiale :

$$P_{fa}(x_i, T) \triangleq \sum_{k=T}^{N_B} \binom{N_B}{k} (p_{fa}(x_i, \varsigma(x_i)))^k (1 - p_{fa}(x_i, \varsigma(x_i)))^{N_B - k}. \quad (3.36)$$

où $p_{fa}(x_i, \varsigma(x_i))$ est la probabilité de fausse alarme au pixel x_i avec le seuil $\varsigma(x_i)$, supposée identique pour tous les pixels $x_j \in B(x_i)$.

Pour autant, la probabilité de fausse alarme (3.36) dépend de $p_{fa}(x_i, \varsigma(x_i))$, mal définie pour l'instant et difficile à estimer en tout point de l'image. Afin de remédier à cela et déterminer $\varsigma(x_i)$, nous avons recours au lemme suivant :

Lemme 3 Soit $V(x_i) \subseteq \mathcal{I}$ un voisinage donné autour du point x_i . D'après l'inégalité de Markov, la probabilité de fausse alarme est bornée de la manière suivante

$$\begin{aligned} \mathbb{P}(S_{N_B}(x_i) \geq T | H_0) &\leq e^{-T} \mathbb{E}[e^{S_{N_B}(x_i)} | H_0] \leq e^{-T} \mathbb{E}[e^{S_{N_B}(x_i)}], \\ &\approx \frac{e^{-T}}{|V(x_i)|} \sum_{x_j \in V(x_i)} e^{S_{N_B}(x_j)}. \end{aligned} \quad (3.37)$$

Néanmoins, le problème de l'estimation des seuils $\varsigma(x_i)$, pour un taux de fausses alarmes donné, n'est pas résolu. La manière la plus conventionnelle consiste à modéliser la densité de probabilité des distances de dissimilarité par un mélange de gaussiennes. Nous avons mis en place un algorithme permettant cela mais il s'est vite avéré que le nombre de modes des densités mesurées est très variable d'une paire d'images à une autre, quelle que soit la taille des motifs ou la taille des voisinages considérés. L'aire des zones occultées joue un rôle prépondérant dans ce constat. Un autre constat intéressant est le fait que les densités de probabilité calculées pour une paire d'images (u_1, u_2) sont très semblables à celles mesurées sur une seule image, c'est-à-dire pour le couple (u_1, u_1) si u_1 est l'image de référence.

3.3.3 Approche non-paramétrique et multi-échelle

Afin de déterminer les seuils de détection $\varsigma(x_i)$, il semble opportun d'examiner une seule image u_1 ainsi que les distances de dissimilarité calculées sur cette image dans des voisinages locaux.

Spécification des distances de dissimilarité

Afin de mesurer la dissimilarité entre motifs (0 lorsque les motifs sont identiques), on peut adopter les mesures suivantes qui ont donné satisfaction dans un grand nombre de cas applicatifs :

$$\phi_{u_1 u_2}(x_i, x_j) = \sum_{\nu \in \mathbb{R}^2} G_a(\nu) (u_1(x_i + \nu) - u_2(x_j + \nu))^2 \quad (3.38)$$

$$\phi_{u_1 u_2}(x_i, x_j) = \sum_{\nu \in \mathbb{R}^2} G_a(\nu) \left(u_1(x_i + \nu) - \frac{u_1^\rho(x_i)}{u_2^\rho(x_j)} u_2(x_j + \nu) \right)^2 \quad (3.39)$$

$$\phi_{u_1 u_2}(x_i, x_j) = 1 - \frac{\sum_{\nu \in \mathbb{R}^2} G_a(\nu) u_1(x_i + \nu) u_2(x_j + \nu)}{\sqrt{\sum_{\nu \in \mathbb{R}^2} G_a(\nu) u_1^2(x_i + \nu)} \sqrt{\sum_{\nu \in \mathbb{R}^2} G_a(\nu) u_2^2(x_j + \nu)}} \quad (3.40)$$

où G_a est un noyau gaussien d'écart-type $a = \sqrt{n/2}$ permettant de donner moins d'importance aux pixels situés en périphérie des motifs considérés et d'introduire une certaine invariance aux rotations de la paire d'images (Buades et al. 2008). La mesure de dissimilarité (3.38) est très classique mais ne permet pas de gérer le cas de changements globaux d'illumination liés à des phénomènes climatiques voire des ombrages ou des spéularités intervenant sur l'une des images. À cet égard, l'hypothèse lambertienne prise en compte dans la mesure de dissimilarité (3.39) permet d'introduire de la robustesse dans la détection (Aach et al. 2001, Brox et al. 2008) où $u_1^\rho = G_\rho \star u_1$ et $u_2^\rho = G_\rho \star u_2$ sont des images obtenues par convolution des images originales u_1 et u_2 par un noyau gaussien d'écart-type ρ . Le paramètre ρ permet de gérer à la fois les changements globaux et locaux d'illumination dans la scène. Enfin, si on s'intéresse aux orientations locales pour détecter les changements, il est recommandé d'utiliser une mesure de corrélation (3.40). D'autres scores sont toujours possibles (Pilet, Strecha & Fua 2008) et les images peuvent être normalisées dans certains cas par égalisation d'histogramme.

Quels seuils de détection ?

Afin de déterminer les seuils de détection, assez problématiques, nous postulons que toutes les décisions positives coïncident avec des distances de dissimilarité plus élevées que la distance la plus élevée au pixel x_i et définie en tout point à partir de l'image de référence u_1 . On définit

$$\zeta(x_i) = \max \left(\sup_{x_j \in b(x_i)} \phi_{u_1 u_1}(x_i, x_j), \bar{\zeta} \right), \quad (3.41)$$

où $b(x_i) = \{x_j \in \mathcal{I} : \|x_j - x_i\| \leq r_b\}$ est un petit voisinage autour du point x_i (e.g. une boule dans \mathbb{R}^2 centrée au point x_i et de rayon r_b). On admet que ce seuil doit être au moins supérieur à une valeur minimale $\bar{\zeta}$ définie comme la moyenne des scores les plus faibles de l'image de référence. Dans cette modélisation, l'arrière-plan est supposé quasi statique mais tolère du mouvement résiduel et peu organisé dont l'origine est liée aux vibrations toujours possibles de la caméra ou à des mouvements naturels du paysage. Toute cette agitation de faible amplitude peut être décrite par des processus stochastiques. Soit $\Delta(x_i) \in \mathbb{R}^2$ un décalage aléatoire (ou translation) tel que $\mathbb{E}[\|\Delta(x_i)\|] = 0$ et $\mathbb{E}[\|\Delta(x_i)\|^2] = \sigma_\Delta^2$. Un motif à la position $x_j \in b(x_i)$ définie comme $x_j = x_i + \Delta(x_i)$ est censé être déplacé à la position x_i dans la seconde image u_2 si aucun changement majeur n'intervient. La taille du voisinage local $b(x_i)$ correspond alors à l'amplitude de la vibration de la caméra ou des dynamiques naturelles en ce point.

Les seuils ainsi définis à partir d'une seule image ne nécessitent pas l'identification de la nature ou du niveau de bruit contrairement aux méthodes usuelles basées sur des tests d'hypothèses paramétriques.

Une approche multi-échelle pour réduire le nombre de fausses alarmes

Afin de mettre en œuvre un algorithme de détection de changement, il est également nécessaire de préciser le seuil T , la taille n des motifs et la taille N_B de la fenêtre de sondage.

Seuil T En posant $T = N_B$, un pixel x_i qui ne trouve aucun correspondant dans la seconde image est déclaré comme pixel occulté. Ce choix assez intuitif a déjà été discuté par Sun et al.

(2005) et Jodoin, Mignotte & Rosenberger (2007) et s'avère efficace dans tous les cas expérimentaux. On suppose qu'un pixel x_i est, à juste raison, considéré comme occulté si toutes les distances de dissimilarité sont plus élevées que $\zeta(x_i)$ pour la fenêtre de sondage $B(x_i)$.

Taille N_B de la fenêtre de sondage Celle-ci dépend de l'amplitude de mouvement considéré comme non-significative pour la détection de changement. En pratique, des fenêtres de taille $N_B = 3 \times 3$ (5×5 au plus) suffisent de manière assez surprenante pour des situations de caméras fixes. En d'autres termes, $|b(x_i)| \equiv |B(x_i)|$ suffit en pratique. Augmenter la taille de la fenêtre de sondage ne devrait pas générer davantage de zones détectées. Néanmoins, il est toujours possible d'apparier des contextes locaux sans réelle relation sur la simple analyse de similarité de motifs locaux.

Représentation multi-échelle des motifs Une approche multi-échelle est généralement recommandée pour analyser des contextes spatiaux avec des détails variés. Dans la modélisation que nous préconisons, une stratégie multi-échelle permet surtout de réduire le nombre de fausses alarmes et d'éviter de fixer précisément la taille p des motifs. Soit D_l une variable aléatoire qui prend la valeur 1 lorsqu'un changement est détecté au pixel x_i et 0 dans le cas contraire, pour une taille de motif $n_l = (2l + 1)^2, 1 \leq l \leq L$. Par la suite, L désigne le nombre de tailles considérée, identique à chaque position, et on considère une collection $\{D_1(x_i), \dots, D_L(x_i)\}$ à chaque pixel x_i .

Les variables binaires de Bernouilli $D_l(x_i)$ sont supposées non identiquement distribuées et pas vraiment indépendantes car se recouvrant bien souvent. Néanmoins, d'après la méthode de Chen-Stein (Arratia, Goldstein & Gordon 1989), $\sum_{l=1}^L D_l(x_i)$ converge en distribution vers une loi de Poisson $Po(\lambda_T(x_i))$ lorsque $L \rightarrow \infty$ avec $\lambda_T(x_i) = \sum_{l=1}^L P_{fa,l}(x_i, T)$ tel que $\mathbb{E}[D_l(x_i)|H_0] = P_{fa,l}(x_i, T), 1 \leq l \leq L$. Par ailleurs, il est également souhaitable de prendre une décision de manière ponctuelle tout en prenant en compte les aspects de contextes spatiaux et multi-échelles analysés au niveau de l'image entière afin de mesurer les événements inhabituels recherchés entre deux images. On peut parvenir à cela assez simplement en considérant des voisinages $V(x_i)$ de grande taille et définis au Lemme 3 servant au calcul de $P_{fa,l}(x_i, N_B)$. Par conséquent, la probabilité de fausse alarme $P_{FA}(x_i, L) \triangleq \mathbb{P}(\sum_{l=1}^L D_l(x_i) > k_D(x_i)|H_0)$ au pixel x_i est donnée par la queue de la loi de Poisson de paramètre $\lambda_{N_B}(x_i) \equiv \lambda_{N_B} = \sum_{l=1}^L P_{fa,l}(N_B)$:

$$P_{FA}(x_i, L) = 1 - \left[\sum_{k=0}^{k_D(x_i)} \frac{e^{-k N_B}}{|\mathcal{I}|^k k!} \left(\sum_{l=1}^L \sum_{x_j \in \mathcal{I}} e^{S_{N_B,l}(x_j)} \right)^k \exp \left(-\frac{e^{-N_B}}{|\mathcal{I}|} \sum_{l=1}^L \sum_{x_j \in \mathcal{I}} e^{S_{N_B,l}(x_j)} \right) \right] \quad (3.42)$$

où $1 \leq (k_D(x_i) + 1) \leq L$ représente le nombre de décisions positives effectivement comptabilisées au pixel x_i pour l'ensemble de tailles de motif considéré. Finalement, on considère qu'un changement a lieu au pixel x_i si $P_{FA}(x_i, L) \leq \alpha(x_i)$ pour un niveau désiré de significativité $\alpha(x_i)$. Sur la base de la stratégie de Bonferroni établie pour les tests multiples, nous fixons $\alpha(x_i) = \epsilon/|\mathcal{I}|$ où $|\mathcal{I}|$ est le nombre de pixels testés et ϵ est le nombre moyen de fausses alarmes désiré.

Algorithme 4 (Détection de changements sur une paire d'images)

Soit L le nombre d'échelles considéré, $\phi_{u_1 u_2}$ une mesure de dissimilarité donnée.

o Pour $l = 1 \dots L$

1. Pour chaque pixel $x_i \in \mathcal{I}$ et des voisinages $b(x_i)$ de taille 3×3 pixels, calculer

$$\varsigma_l(x_i) = \max \left(\sup_{x_j \in b(x_i)} \phi_{u_1 u_2, l}(x_i, x_j), \bar{\varsigma}_l \right).$$

2. Pour chaque pixel $x_i \in \mathcal{I}$ et des fenêtres de sondage $B(x_i)$ de taille 3×3 pixels $B(x_i)$, calculer $S_{N_B, l}(x_i) = \sum_{x_j \in B(x_i)} \mathbb{1}[\phi_{u_1 u_2, l}(x_i, x_j) \geq \varsigma_l(x_i)]$.

3. Calculer $P_{fa, l}(N_B) = \frac{\epsilon^{-N_B}}{|\mathcal{I}|} \sum_{x_j \in \mathcal{I}} e^{S_{N_B, l}(x_j)}$.

o Calculer $\lambda_{N_B} = \sum_{l=1}^L P_{fa, l}(N_B)$.

o Pour chaque pixel $x_i \in \mathcal{I}$, prendre la décision $H_\epsilon(u_1, u_2)(x_i) = \mathbb{1}[P_{FA}(x_i, L) \leq \epsilon/|\mathcal{I}|]$ où

$$P_{FA}(x_i, L) = 1 - \sum_{k=0}^{k_D(x_i)} (\lambda_{N_B})^k \frac{e^{-\lambda_{N_B}}}{k!}.$$

Propriétés de l'algorithme L'approche proposée repose sur les grands principes de la méthode *a contrario* proposée par Desolneux et al. (2000) permettant un contrôle du nombre fausses alarmes analogue à la stratégie de Bonferroni mais autorisant des valeurs supérieures à 1. En pratique, on fixe $\epsilon = 1$, ce qui sous-entend qu'en moyenne un pixel est faussement déclaré comme occulté mais les autres pixels sont « significatifs » (Desolneux et al. 2000).

Soit $H_\epsilon(u_1, u_2) : \mathcal{I} \rightarrow \{0, 1\}$ la carte de détection telle que : $H_\epsilon(u_1, u_2)(x_i) = 1$ si $P_{FA}(x_i, L) \leq \epsilon/|\mathcal{I}|$ et 0 sinon. L'algorithme est alors très simple et est présenté ci-après. Mis à part le fait que nul n'est besoin de modéliser très précisément l'image et le bruit, l'approche est invariante aux changements de contraste linéaires, aux translations et rotations globales appliquées aux deux images. La figure 3.6 présente quelques résultats sur des paires d'images pour lesquelles les vérités-terrain ont été établies (Toyama, Krumm, Brumitt & Meyers 1999, Heikkila & Pietikainen 2006). Nous présentons également un résultat de détection de changements dans le cadre applicatif de la restauration de « vieux films » (Fig. 3.7).

À nouveau, le point de vue bayésien Au lieu de seuiller la probabilité de fausse alarme $P_{FA}(x_i, L)$, il peut-être intéressant de reconsidérer une modélisation par champs aléatoires de la paire d'images. Soit \mathcal{Z} l'ensemble des configurations $\mathbf{z} = (z(x_i))_{x_i \in \mathcal{I}}$ avec $z(x_i) \in \{0, 1\}$ pour tout $x_i \in \mathcal{I}, i = 1, \dots, N$. Dans le cadre des champs aléatoires conditionnels et discriminants (Kervrann et al. 2009) nous pouvons définir une fonctionnelle globale d'énergie que l'on cherche à minimiser par un algorithme « *min-cut/max-flow* » (Greig et al. 1989, Boykov et al. 2001) de manière rapide et efficace :

$$E_{u_1 u_2}^{count}(\mathbf{z}) = \sum_{x_i \in \mathcal{I}} \sum_{l=1}^L a_l \left| \frac{1}{N_B} S_{N_B, l}(x_i) - z(x_i) \right| - \theta \sum_{\langle x_i, x_j \rangle} \mathbb{1}[z(x_i) \neq z(x_j)] \quad (3.43)$$

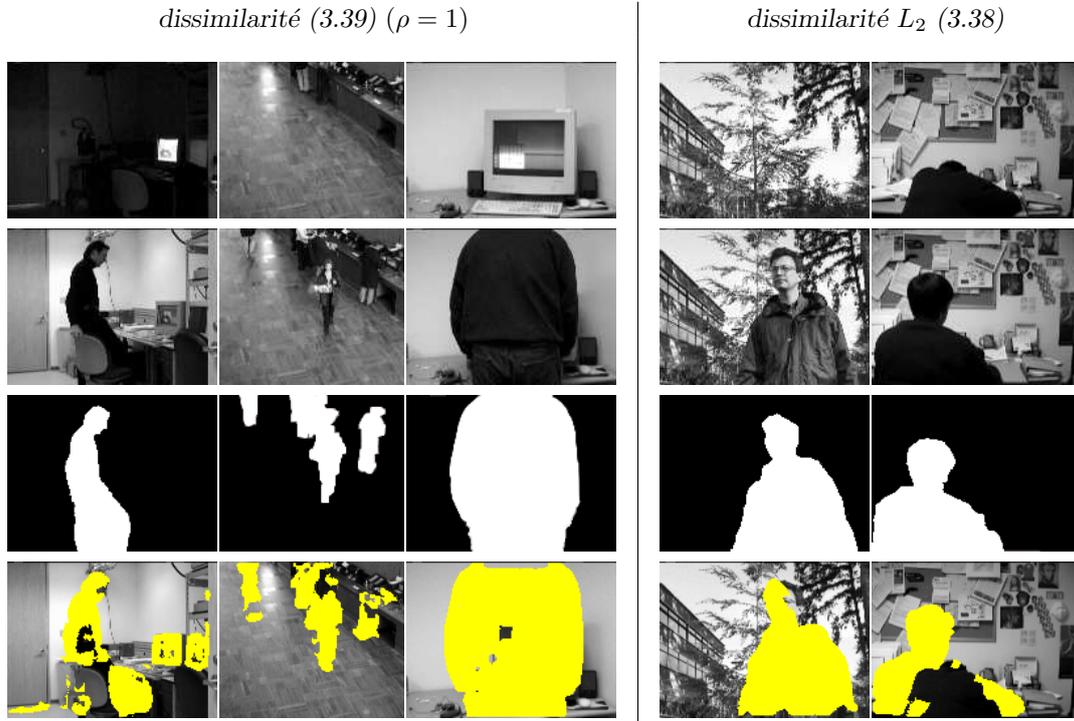


FIG. 3.6. Résultats expérimentaux. Deux premières lignes : paires d'images ; troisième ligne : vérités terrain ; quatrième ligne : détection de changements.

et ceci dans l'esprit des méthodes classiques de détection de changements (Jodoin, Mignotte & Rosenberger 2007). Le premier terme de cette énergie cherche à labéliser un pixel avec une étiquette « changement » si le nombre de décisions locales positives est élevé et proche de N_B , ceci pour toutes les tailles de motifs considérées où $a_l = \frac{a'_l}{\sum_{l=1}^L a'_l}$ avec $a'_l = (2l + 1)^{-1}$ sont des poids qui donnent davantage d'importance aux motifs de petite taille. Le second terme est relativement usuel puisqu'il s'agit du modèle d'Ising (structure de voisinage « 4-voisins ») préconisé pour la modélisation d'images binaires dans le cadre markovien, où $\theta > 0$ est un paramètre qui permet d'équilibrer les deux termes d'énergie. On impose davantage de régularité à l'image estimée en assignant une forte valeur à θ . Dans ce cas, le nombre de composantes connexes et leurs périmètres ont tendance à baisser. Généralement, il n'est pas très aisé d'obtenir des résultats équivalents même en modifiant θ . Notons que nous pouvons contrôler la performance du détecteur, sur la base de calculs de probabilités de fausse alarme menés en tout point de l'image, contrairement aux estimateurs MAP (maximum *a posteriori*).

3.4 Conclusion et perspectives

Dans ce chapitre, nous avons exposé en détails quelques contributions à l'analyse d'images par des méthodes non-paramétriques (Kervrann et al. 2009, Kervrann et al. 2009). Nous avons

attiré l'attention sur deux problèmes fondamentaux : le débruitage d'images et la détection de changements dans une paire d'images. Une littérature très abondante existe sur ces deux sujets. Nous avons privilégié les approches non-paramétriques de manière à ne pas imposer des modèles d'images, souvent incomplets et pas toujours capables de saisir toute la complexité des images. Les approches que nous défendons ont également un intérêt dès qu'il s'agit d'évaluer la fiabilité des solutions calculées. Dans le premier cas, nous avons cherché un estimateur local contrôlé par un « oracle ». Dans le second cas, nous nous sommes intéressés aux probabilités de fausse alarme afin de caractériser un détecteur. Ces points de vue, vraisemblablement un peu austères, ont néanmoins conduit à des algorithmes originaux et assez efficaces si on les compare à l'état de l'art sur ces deux sujets.

Par ailleurs, nous avons opté pour ce cadre de modélisation non-paramétrique afin de répondre aux nombreuses questions de nos collaborateurs de biologie cellulaire qui exploitent les images en microscopie. Les réflexions que je viens d'exposer ont en effet été conjointement menées en partenariat avec des laboratoires de biologie fondamentale pour lesquels la microscopie est un instrument technologique d'investigation. L'image permet d'analyser des mécanismes spatiaux et temporels qu'il s'agit de décrypter, à l'échelle d'une cellule vivante. Par conséquent, il est plus prudent d'analyser les informations sans « trop » d'*a priori*, afin de pouvoir avancer des thèses objectives.

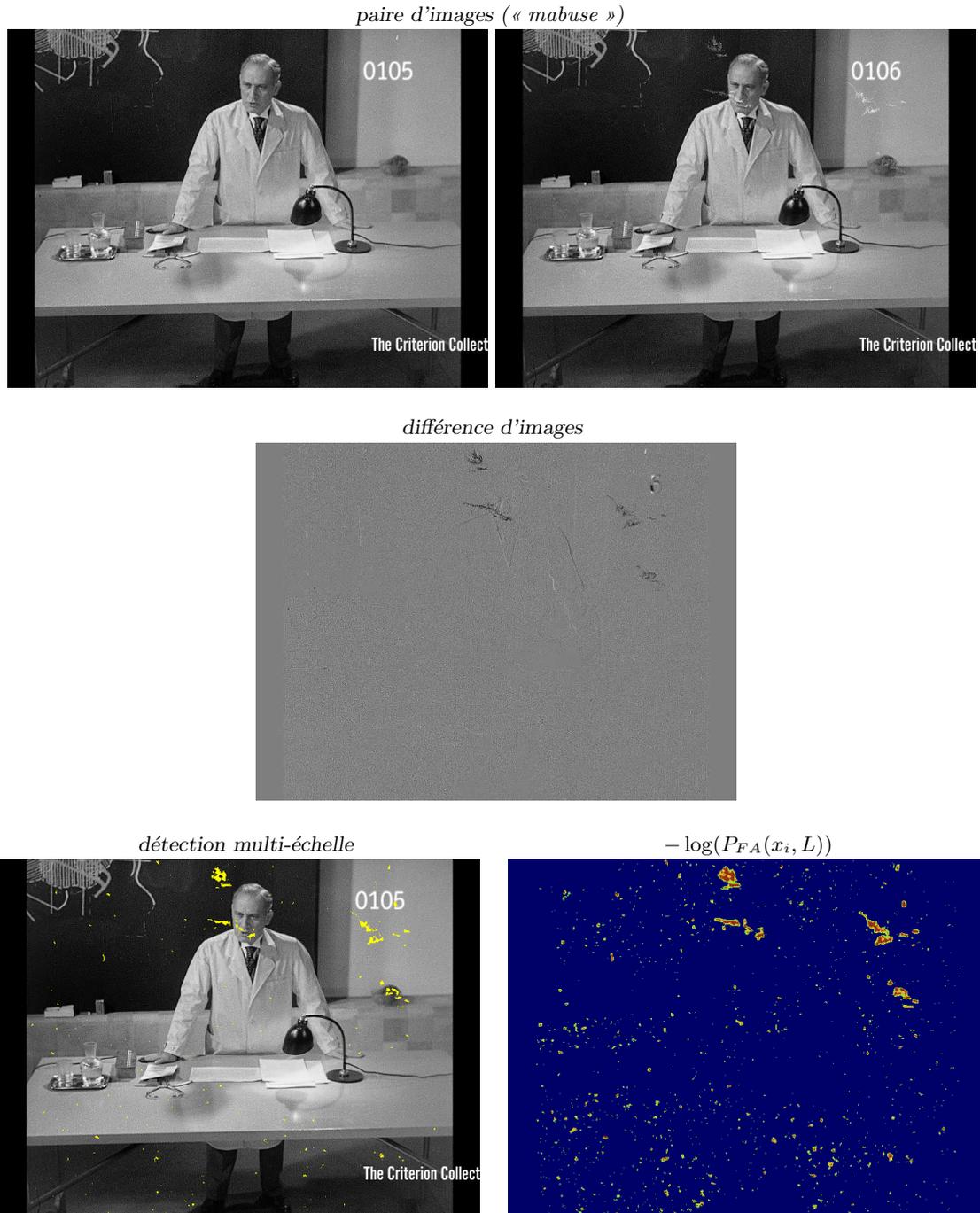


FIG. 3.7. Détection de défauts sur une paire d'images extraite d'un « vieux film ».

Modèles et algorithmes en action : analyse du transport intra-cellulaire

En biologie, comme dans pratiquement toutes les disciplines en sciences, les approches scientifiques sont à la fois poussées et limitées par les technologies. Ces dernières années, parallèlement aux techniques et aux concepts de la génomique à haut débit, qui est l'un des moteurs de la recherche en biologie, les avancées en microscopie multi-dimensionnelle couplées aux nouvelles techniques de marquage par sondes fluorescentes (« *Green Fluorescent Protein* ») (Shimomura, Johnson & Saiga 1962, Chalfie, Tu, Euskirchen, Ward & Prasher 1994) ont révolutionné la biologie moléculaire et cellulaire (Salamero 2008). L'imagerie photonique (les cellules expriment une protéine de fusion fluorescente en réponse à une excitation lumineuse (e.g. laser)) dite à haute résolution spatiale et temporelle joue désormais un rôle essentiel pour sonder les processus moléculaires des interactions des protéines (expression et fonction des gènes) dans différents compartiments ou domaines cellulaires. Des avancées considérables en biologie fondamentale ont ainsi déjà été effectuées dans la description des grands principes qui régulent le fonctionnement cellulaire, les principales voies de transport, les mécanismes de tri et d'adressage des protéines, dans le maintien ou la défaillance de l'intégrité fonctionnelle des compartiments de la cellule. Il est également établi qu'un nombre restreint de protéines interagissent fonctionnellement et dynamiquement dans des sous-domaines de chaque compartiment. Ces interactions sont nécessaires pour assurer des fonctions vitales, à des échelles spatiales et temporelles variées sur l'ensemble du cycle cellulaire, que les cellules soient en division ou en interphase.

Acquérir une vue quantifiée plus complète de la physiologie de la cellule pour toutes les échelles d'observation spatiales « nano-micro » et temporelles, est le défi qu'il convient de relever désormais. Mais force est de constater que l'analyse des données de microscopie collectées, ne serait-ce que ces trois dernières années, via différentes modalités de microscopie optique (modalités d'imagerie¹ WF, MSDC, FLIM-FRET, TIRF, SIM...) et électronique (TEM), est très problématique.

¹voir glossaire

Dans une expérience biologique, il s'agit en effet de traiter des quantités considérables de volumes 3D, 2D+T ou 3D+T voire nD+T (imagerie multi-spectrale), pouvant représenter plusieurs centaines de méga-octets (Pepperkok & Ellenberg 2006, Neumann, Held, Liebel, Erfle, Rogers, Pepperkok & Ellenberg 2006). Leur contenu est relativement nouveau et original, mais leur traitement avec des algorithmes existants de traitement d'images, est souvent peu efficace, voire mal adapté. L'information recherchée est parfois peu prévisible puisque l'instrument de microscopie est précisément utilisé pour mettre en évidence un processus, une interaction ou un mécanisme moléculaire inattendu ou rarement observé. Notons que l'analyse visuelle de ces données-images devient presque impossible puisqu'il s'agit d'inspecter des centaines de séquences temporelles d'images volumiques, puis de les manipuler lors du passage d'une échelle spatiale à l'autre, par exemple dans le cadre d'imagerie(s) corrélative(s) (« *Correlative Light and Electron Microscopy* ») (Vicidomini, Gagliani, Canfora, Cortese, Frosi, Santangelo, Di Fiore, Boccacci, Diaspro & Tacchetti 2008). Pourtant, toutes les informations partielles ou descripteurs extraits automatiquement doivent être intégrés rapidement dans des modèles biologiques ou biophysiques à des fins de prédiction, pour des applications dans les domaines des biotechnologies et de la santé. Dans un tel contexte, nous sommes confrontés à la fois aux problèmes de gestion de « masses de données », d'estimation en « grande dimension », de modélisation de « systèmes complexes » et de « changements d'échelles ».

Dans un tel contexte, il est nécessaire d'entretenir un dialogue entre expérimentation, observation, et modélisation, malgré le cloisonnement disciplinaire souvent contraignant. Ces dernières années, j'ai essentiellement porté mes efforts sur la compréhension de quelques mécanismes régulant le ciblage et le transport « membranaires » de protéines dans la cellule, mécanismes observés grâce des instruments de microscopie optique très performants et disponibles à l'Institut Curie (Umr 144 Cnrs Institut Curie, PICT-IBiSA (Plateforme Imagerie Tissulaire et Cellulaire), « *Nikon Imaging Center* »²) (Fig. 4.1). L'étude des mécanismes en question est une problématique très générique et représente l'un des domaines de recherche les plus actifs en biologie cellulaire. Les investigations dans ce domaine doivent en effet permettre de quantifier à moyen terme le transport intra-cellulaire et les causes de son éventuelle détérioration.

4.1 La vidéo-microscopie : de nouveaux défis en analyse d'images

La vidéo-microscopie de fluorescence, principalement utilisée dans mes activités de recherche joue donc un rôle majeur et se singularise surtout par la nature des objets observés (organites de la cellule), par un nombre souvent élevé d'éléments mobiles de petite taille et éventuellement coalescents (fusion), par la complexité des dynamiques mettant en jeu plusieurs entités ou groupes d'entités en interaction. Face à ce type d'observations, les méthodes de filtrage stochastique pour le suivi d'objets (méthodes « particulières ») (Gordon, Salmond & Smith 1993) ne sont pas toujours capables de traiter des données aussi complexes si elles ne sont adaptées au contexte applicatif (Genovesio, Liedl, Emiliani, Parak, Coppey-Moisan & Olivo-Marin 2006, Smal, Draegestein, Galjart, Niessen & Meijering 2008, Smal, Meijering, Draegestein, Galjart, Grigoriev, Akhmanova, van Royen, Houtsmuller & Niessen 2008, Chenouard, Bloch & Olivo-Marin 2009). Les méthodes d'analyse de mouvement analytique sont également peu adaptées à des situations

²<http://nimce.curie.fr>

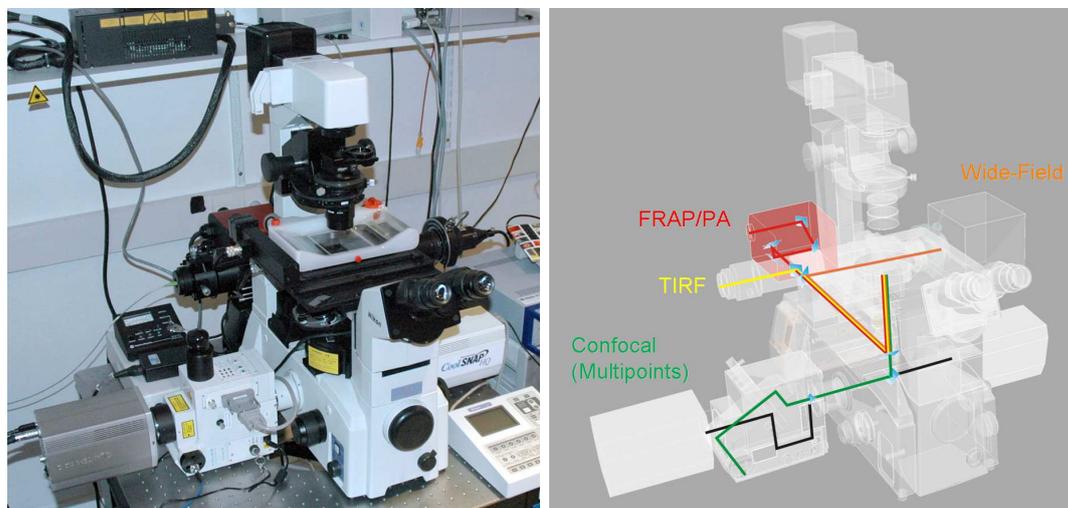


FIG. 4.1. Montage et instrument de microscopie multi-modale de la PICT-IBiSA (Umr 144 Cnrs Institut Curie) / « Nikon Imaging Center ».

de coalescence d'objets, mal résolus du point de vue de l'optique, et à des images au niveau de bruit élevé. Par ailleurs, les outils de traitement d'images mis à la disposition des biologistes et des microscopistes permettent seulement de manipuler et de visualiser le volume brut d'information, puis d'exécuter des opérations semi-automatiques et très sommaires d'extraction de contours des objets d'intérêt. La charge de travail qui consiste *in fine* à effectuer des mesures morphométriques robustes sur les objets étudiés (surface, volume, forme, nombre, organisation spatiale ...) paraît vite bien lourde à la vue des outils disponibles. De surcroît, l'information 3D, si chèrement acquise, est souvent résumée en définitive à une information 2D (après une opération de moyennage), plus facile à maîtriser par les utilisateurs. Notre objectif est donc de concevoir des modèles et des algorithmes originaux de traitement, aptes à gérer le volume et la complexité des informations véhiculées par ces images de microscopie 2D+T, 3D et 3D+T.

Une première contribution significative visait à développer une méthode incrémentale originale de correction de l'atténuation du signal de fluorescence selon la direction axiale (profondeur de l'échantillon). Celle-ci s'appuyait sur des techniques classiques de régression robuste (M-estimateurs) couplées à une modélisation de l'image issue des travaux menés en vision par ordinateur pour calculer le flot optique (Odohez & Bouthemy 1995, Black & Rangarajan 1996b, Black & Anandan 1996, Negahdaripour 1998, Haussecker & Fleet 2001). En soi, l'approche n'avait pas un caractère très original mais l'algorithme proposé, simple dans son usage, s'est avéré bien adapté pour résoudre un problème classique en microscopie optique (fluorescence). Il fut conçu dans le cadre de la thèse de L. Pardini³ (2002) dirigée par A. Trubuil et B. Kaeffer) et l'algorithme a fait l'objet d'une publication dans une revue internationale de microscopie (Kervrann,

³Inra-UFNDH Nantes / « Analyse multi-dimensionnelle de bio-marqueurs le long de l'axe de la crypte colique en fonction du régime alimentaire »

Legland & Pardini 2004). Dans ce chapitre, je présente une partie des travaux qui ont suivi. Ils ont trait au développement de méthodes statistiques et d'algorithmes issus de travaux récents en vision par ordinateur, en traitement du signal, en statistiques non-paramétriques et en théorie des graphes, afin d'améliorer, d'extraire et d'organiser les informations dynamiques relatives à des séries temporelles d'images obtenues en vidéo-microscopie de fluorescence. À chaque occasion, nous avons proposé des méthodes innovantes pour répondre à ces objectifs, méthodes adaptées au contexte de la microscopie *in vivo*. La thématique que nous explorons offre par ailleurs l'avantage de proposer des problématiques nouvelles (e.g. déplacement de centaines de particules/spots sur réseaux complexes) pouvant susciter des modélisations et méthodes d'analyse particulières qui ne manqueront pas d'être appliquées à d'autres modalités d'imagerie. Ces actions sont menées de manière coordonnée, avec prise en compte régulière des résultats partiels obtenus par chaque partenaire. Cette « biologie numérique » nécessite la révision constante, car par définition itérative, des méthodologies de biologie, d'instrumentation et d'informatique, de traitement d'images mises en œuvre, ceci afin d'améliorer les représentations, de les complexifier, de les simplifier ou d'en créer de nouvelles plus adaptées.

4.2 Débruitage spatio-temporel en vidéo-microscopie

Le principe de la méthode de débruitage a été exposé dans le chapitre 3 et a été décrit en détail(s) dans (Boulanger et al. 2010, Boulanger et al. 2008, Boulanger et al. 2007). Dans ce paragraphe, nous décrivons surtout la méthode de « stabilisation de variance » qui permet notamment de traiter le cas du bruit de Poisson, principal responsable du faible rapport signal-sur-bruit mesuré sur les images de vidéo-microscopie.

En microscopie « champ-large » (WF) ou « confocale rapide » (MSDC), le faible nombre de photons collectés au niveau du capteur et l'agitation thermique induite par les composants électroniques doivent être pris en compte dans le processus de formation de l'image. On considère que le bruit mesuré est du type gaussien-Poisson et le modèle statistique d'images à considérer est de la forme suivante :

$$Z(x_i, t) = \mathcal{G}_0 \aleph(x_i, t) + e(x_i, t), \quad i \in \{1, \dots, N\}, \quad t \in \{1, \dots, T\} \quad (4.1)$$

où $Z(x_i, t)$ désigne l'intensité observée à la position $x_i \in \mathcal{I} \subset \mathbb{R}^d$ ($d = 2, 3$) et à l'instant t dans l'image volume spatio-temporel, et \mathcal{G}_0 représente le gain global du système électronique. Le nombre $\aleph(x_i, t)$ de photons est une variable aléatoire qui suit une loi de Poisson de paramètre $\kappa(x_i, t)$:

$$p(\aleph(x_i, t)) = \frac{\kappa(x_i, t)^{\aleph(x_i, t)} e^{-\kappa(x_i, t)}}{\aleph(x_i, t)!}. \quad (4.2)$$

Dans ce modèle relativement simple, on assimile le « courant d'obscurité » à un bruit blanc gaussien de moyenne non-nulle $\mathbb{E}[e(x_i, t)] = \nu_e$ et de variance $\text{Var}[e(x_i, t)] = \sigma_e^2$. Les variables aléatoires $\aleph(x_i, t)$ et $e(x_i, t)$ sont supposées indépendantes. L'objectif est d'estimer la variable $\aleph(x_i, t) = \mathcal{G}_0 \kappa(x_i, t) + \nu_e$ à la position x_i à l'instant t à partir des données bruitées $Z(x_i, t)$.

La stratégie que nous avons privilégiée pour traiter cette situation consiste à stabiliser la

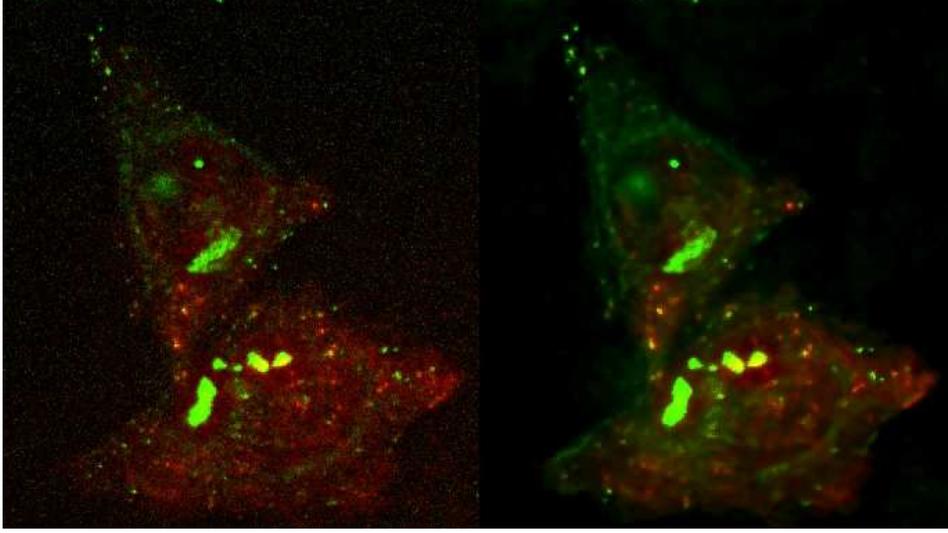


Fig. 4.2. GFP-Rab6A (cellule HeLa) : débruitage d'une séquence 3D+T en microscopie « confocale rapide » MSDC (vert) (projection de 10 plans selon l'axe axial) et d'une séquence 2D+T en microscopie TIRF (rouge) (modalités "nativement" recalées).

variance du bruit selon la démarche initialement proposée il y a quelques années par Murtagh, Starck & Bijaoui (1995). La « Transformée d'Anscombe Généralisée » est de la forme :

$$\mathcal{T}_{AG}(Z(x_i, t)) = \frac{2}{\mathcal{G}_0} \sqrt{\mathcal{G}_0 Z(x_i, t) + \frac{3}{8} \mathcal{G}_0^2 + \sigma_e^2 - \mathcal{G}_0 \nu_e} \quad (4.3)$$

et permet en théorie de « convertir » un bruit de type Poisson en un bruit gaussien de variance unitaire : $\text{Var}[\mathcal{T}_{AG}(Z(x_i, t))] = 1$. Cette transformée nécessite néanmoins l'estimation des paramètres \mathcal{G}_0 , σ_e^2 et ν_e du modèle de formation de l'image. Un premier algorithme a été proposé par Starck & Murtagh (1998) pour estimer itérativement ces paramètres. À l'instar de Delpretti, Luisier, Ramani, Blu & Unser (2008) et Ramani, Vonesch & Unser (2008), nous avons préféré exhiber le modèle linéaire suivant :

$$\text{Var}[Z(x_i, t)] = \mathcal{G}_0 \mathbb{E}[Z(x_i, t)] + \sigma_e^2 - \mathcal{G}_0 \nu_e. \quad (4.4)$$

Une régression dans le plan $(\mathbb{E}[Z(x_i, t)], \text{Var}[Z(x_i, t)])$ permet en effet d'identifier les paramètres \mathcal{G}_0 et la variable $\Delta_{CO} = \sigma_e^2 - \mathcal{G}_0 \nu_e$, suffisante dans notre cas pour stabiliser la variance. Sur le plan pratique, les moyennes et les variances empiriques du couple $(\mathbb{E}[Z(x_i, t)], \text{Var}[Z(x_i, t)])$ sont préalablement estimées (M-estimateurs) sur des blocs de pixels 4D de taille variable selon l'homogénéité mesurée sur chaque bloc.

Dans l'étape suivante de la procédure d'estimation, nous considérons l'algorithme (3.2.2)

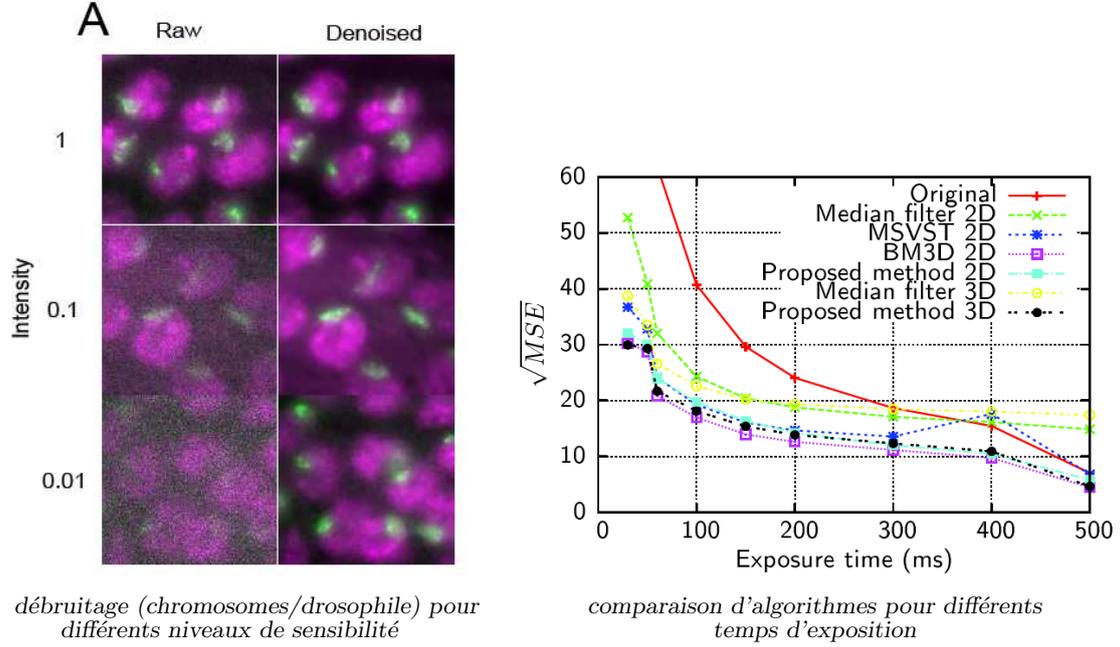


FIG. 4.3. Analyse des performances du débruitage pour différents temps d'exposition (droite) et niveaux de sensibilité du laser (gauche).

adapté au traitement du bruit additif gaussien (Section 3.2) :

$$\mathcal{T}_{AG}(Z(x_i, t)) = u(x_i, t) + \varepsilon(x_i, t) \quad \varepsilon(x_i, t) \sim \mathcal{N}(0, 1). \quad (4.5)$$

où $x_i \in \mathcal{I} \subset \mathbb{R}^d$ ($d = 2, 3$). Cet algorithme de débruitage permet d'estimer une image \hat{u} à laquelle on applique une « Transformée d'Anscombe Généralisée Inverse » : $\hat{\mathcal{Z}}(x_i, t) = \mathcal{T}_{GA}^{-1}(\hat{u}(x_i, t))$. Sur le plan expérimental, nous avons enfin vérifié que l'estimateur ainsi défini, est biaisé et tend vers $1/4$ si le nombre de photons est suffisamment élevé. Au final, un estimateur non-biaisé est obtenu en corrigeant empiriquement l'estimation précédente à l'aide de simulations. Nous obtenons finalement :

$$\hat{\mathcal{Z}}_{NB}(x_i, t) = \mathcal{T}_{GA}^{-1}(\hat{u}(x_i, t)) + \frac{1}{4} (1 - \exp(-1, 3 \mathcal{T}_{GA}^{-1}(\hat{u}(x_i, t)) + C_A)) \quad (4.6)$$

$x_i \in \mathcal{I} \subset \mathbb{R}^d$ et $t \in \{1, \dots, T\}$. Contrairement aux idées préconçues, la stabilisation de variance ainsi corrigée permet d'obtenir des résultats tout à fait satisfaisants dès que $\kappa(x_i, t) > 3$ et non $\kappa(x_i, t) > 30$.

Ce travail a été décrit et validé principalement dans (Boulanger et al. 2010) et la méthode a été évaluée expérimentalement très récemment sur un grand nombre de séquences d'images (Carlton, Boulanger, Kervrann, Sibarita, Salamero, Gordon-Messer, Haber, Haase, Shao, Winoto, Matsuda, Kner, Usawa, Strukov, Gustafsson, Kam, Agard & Sedat 2010). La figure 4.4 illustre

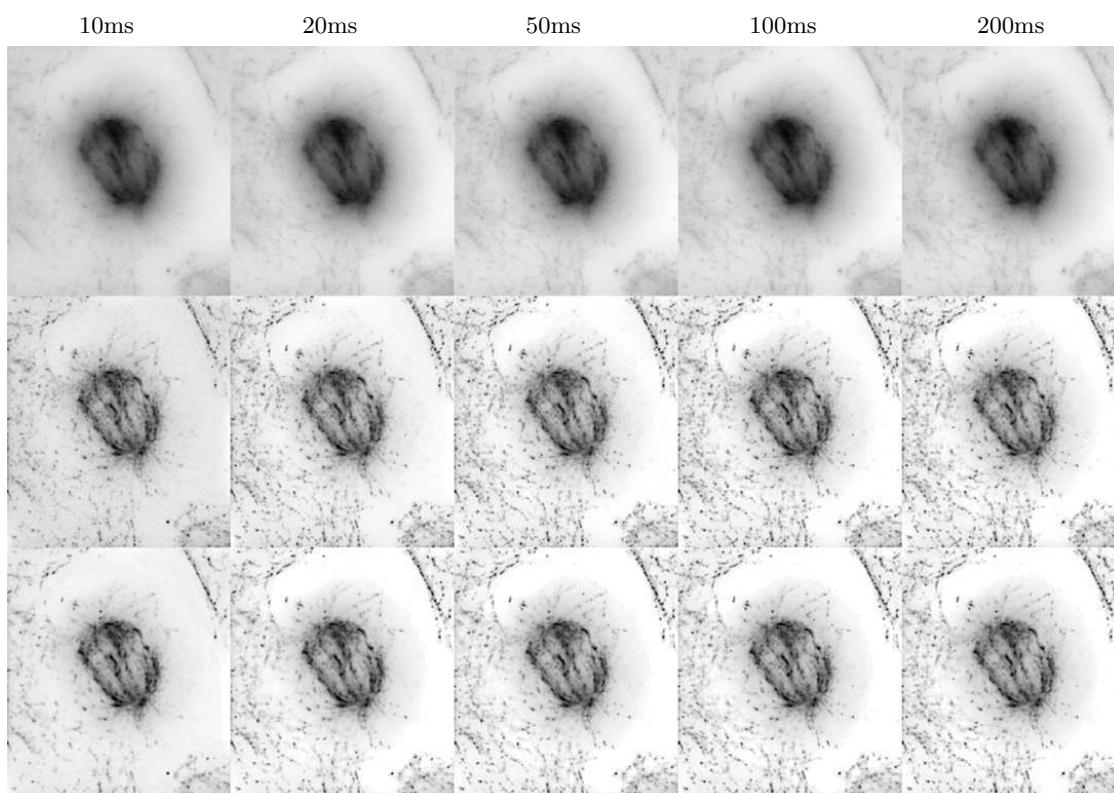


FIG. 4.4. Cellule HeLa fixée observée avec cinq temps d'exposition différents (image 3D $200 \times 200 \times 36$ pixels). La première ligne présente la projection d'intensité maximum des 36 coupes selon la direction axiale et les deuxième et troisième lignes présentent les résultats de déconvolution avec l'algorithme Gold-Meinel (Sibarita et al. 2002) et ceux obtenus en débruitant le volume préalablement avec l'algorithme 'ND-SAFIR'.

les performances de l'algorithme lorsque celui-ci précède un algorithme rapide de déconvolution (Gold-Meinel) (Sibarita, Magnin & De Mey 2002). Nous avons également mis en compétition plusieurs algorithmes performants de débruitage (Zhang, Fadili & Starck 2006, Dabov et al. 2007) pour évaluer leurs efficacités respectives (Figs. 4.5 et 4.3). Le logiciel 'ND-SAFIR'⁴ a enfin été mis à disposition auprès de plusieurs laboratoires académiques et instituts dont les thématiques relevaient principalement de la biologie cellulaire et de la biophysique : Inra, Institut Curie, IMCB Singapour, UCSF San-Francisco, Friedrich Miescher Institute, of Basel, University of Oxford, University of Heriot Watt, Ludwig Maximilians University, Munich, Cold Spring Harbor Laboratory, Weizmann Institute, University of Jerusalem, University of Newcastle, University of Tampere, Fred Hutchinson Cancer Research Center ...

⁴(C++, JAVA, Matlab / Linux, Windows and MacIntosh OS)

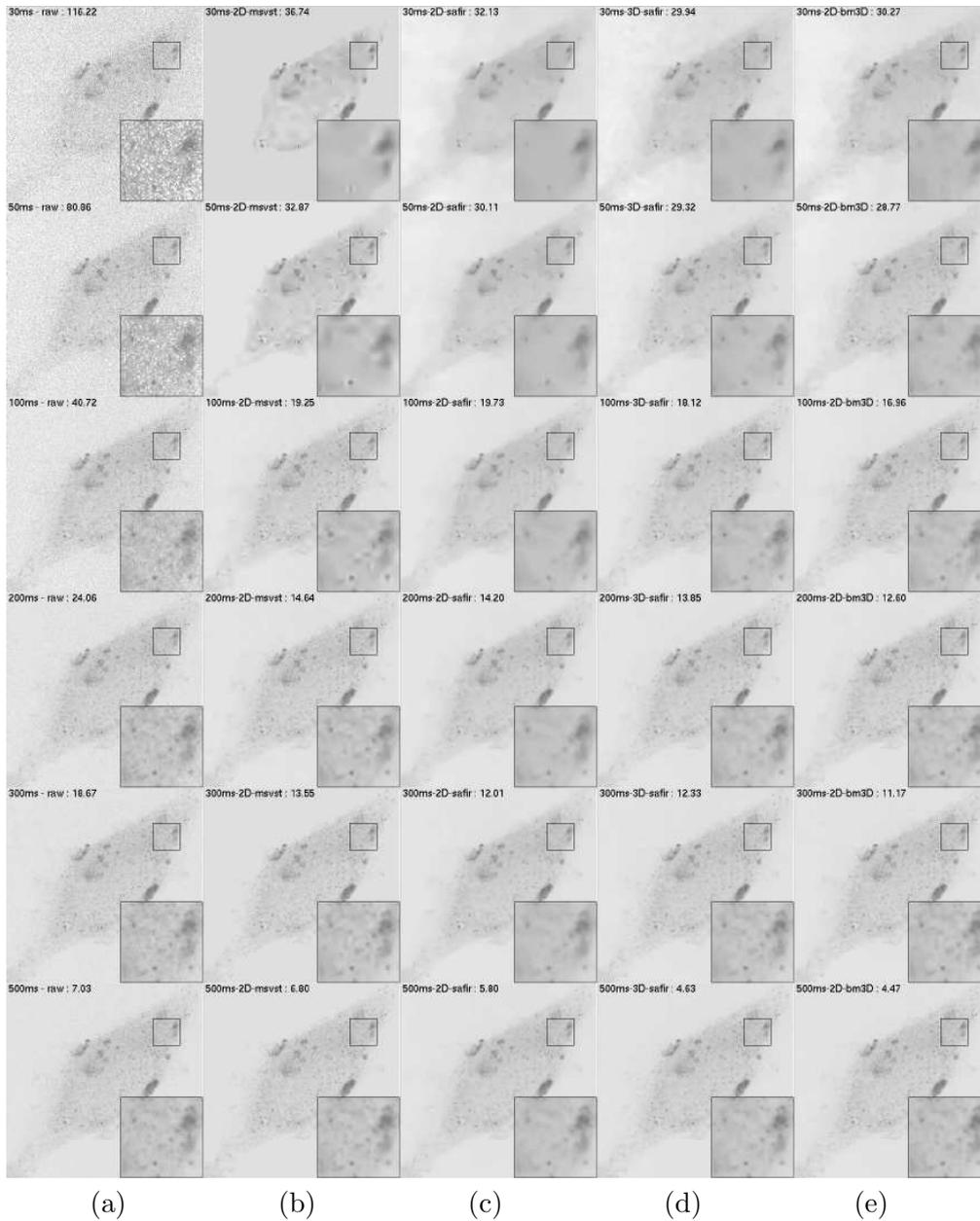


FIG. 4.5. Expériences sur cellule fixée HeLa marquée par GFP-Rab6 et observée en microscopie « confocale rapide » (MSDC). La première colonne présente une coupe 2D d'un volume original 3D pour des temps d'exposition compris entre 30ms et 500ms. De gauche à droite, les colonnes présentent respectivement les résultats obtenus avec la méthode de stabilisation de variance multi-échelle (Zhang et al. 2006) (b), l'algorithme 'ND-SAFIR' en traitant chaque coupe 2D indépendamment (c), l'algorithme 'ND-SAFIR' en considérant le volume 3D (d) et enfin l'algorithme très performant BM3D (Dabov et al. 2007) (e). Les chiffres indiqués correspondent aux erreurs quadratiques moyennes en prenant comme image de référence, l'image acquise avec un temps d'exposition de 500ms (dernière ligne).

4.3 Correction du photo-blanchiment et soustraction de « fond »

En microscopie « champ-large », nous sommes vite confrontés au problème de photo-blanchiment qui s'apparente à une extinction progressive et lente de la sonde fluorescente (Garcia-Parajo, Segers-Nolten, Veerman, Greve & van Hulst 2000), abaissant le rapport signal-sur-bruit au cours du temps. Ce phénomène est très visible si on analyse les variations d'intensité entre deux volumes 3D consécutifs acquis en début de séquence. Afin de le corriger, on impose bien souvent que la moyenne d'intensité à l'intérieur de la cellule reste constante et qu'un modèle exponentiel est suffisamment précis pour décrire l'évolution temporelle de la fluorescence. On considère le modèle paramétrique suivant :

$$\bar{I}_{cell}(t) - \bar{I}_{fond}(t) = A_1 e^{-t/\mathcal{T}_b} + A_2 \quad (4.7)$$

où t désigne la variable temporelle, $\bar{I}_{cell}(t)$ et $\bar{I}_{fond}(t)$ désignent la moyenne de l'intensité respectivement à l'intérieur et l'extérieur de la cellule. Les constantes A_1 and A_2 sont les paramètres du modèle et \mathcal{T}_b est une constante de temps qui contrôle le déclin de la fluorescence. Notons que, idéalement, la constante A_2 est nulle, ce qui est rarement le cas en pratique.

Dans (Boulanger, Gidon, Kervram & Salamero 2009), nous avons proposé de corriger le photo-blanchiment à partir des paramètres estimés à l'aide d'une procédure de minimisation Gauss-Newton. L'intensité corrigée $Y(x_i, t)$ est de la forme suivante,

$$Y(x_i, t) = (f(x_i, t) - \bar{I}_{fond}(t) - A_2) e^{t/\mathcal{T}_b}. \quad (4.8)$$

où $f(x_i, t)$ est l'intensité mesurée au pixel $x_i \in \mathcal{I} \subset \mathbb{R}^d$ et à l'instant $t \in \{1, \dots, T\}$.

Dans certains cas (e.g. microscopie WF), il est souhaitable d'effectuer une stabilisation de variance suivie d'une correction du photo-blanchiment, ce qui revient à calculer à chaque position spatiale x_i :

$$Y(x_i, t) = (T_{GA}(f(x_i, t)) - T_{GA}(\bar{I}_{fond}(t))) e^{t/2\mathcal{T}_b}. \quad (4.9)$$

La variance du bruit est alors multipliée par un facteur $e^{t/2\mathcal{T}_b}$.

4.3.1 Approche spatio-temporelle et semi-paramétrique

D'une manière plus générale, nous avons cherché à identifier les composantes d'un signal temporel extrait d'une séquence d'images. Soit le signal monodimensionnel mesuré à la position $x_i \in \mathcal{I}, i = 1, \dots, N$, et représenté par un vecteur $(f(x_i, 1), \dots, f(x_i, T))^T$ de dimension T . On considère le modèle suivant pour décrire les éléments de ce vecteur :

$$f(x_i, t) = v(x_i, t) + c(x_i, t) + \xi(x_i, t), \quad t \in \{1, \dots, T\} \quad (4.10)$$

où $\xi(x_i, t)$ représente les erreurs de modélisation assimilées à un bruit blanc gaussien. La composante $v(x_i, t) \geq 0$ permet de décrire les « objets » qui se superposent au « fond » $c(x_i, t)$, aux variations temporelles « lentes ».

Plusieurs hypothèses sont avancées pour expliquer l'hétérogénéité des vitesses de décroissance de l'intensité dans le « fond » : 1/ dans certaines régions, la haute concentration de fluorophores

peut conduire à une surexposition lumineuse due à la fluorescence de molécules avoisinantes ; 2/ des pigments présents dans certains tissus peuvent absorber les photons incidents et ainsi conduire à une vitesse de décroissance plus faible ; 3/ la concentration de réactifs peut varier localement. En particulier la teneur en oxygène peut atteindre des concentrations élevées dans certains lipides. Il existe d'autres phénomènes à l'origine de l'évolution de l'intensité du « fond » (Hoogenboom, van Dijk, Hernando, van Hulst & Garcia-Parajo 2005). Ainsi la diffusion de molécules fluorescentes dans le cytosol ou dans des membranes décrites par des modèles de réaction-diffusion est également responsable de l'évolution de l'intensité. L'importance relative de chacun de ces phénomènes dépend du type d'expérience et d'un nombre important de paramètres. Il est très probable que les courbes d'évolution observées soient alors le résultat de la combinaison de plusieurs phénomènes. Nous avons ainsi été amenés à considérer trois modèles paramétriques pour décrire l'évolution temporelle de la composante « fond » :

$$\begin{cases} c(x_i, t) = a_1(x_i) + a_2(x_i) t & (M1) \\ c(x_i, t) = a_1(x_i) + a_3(x_i) e^{-a_4(x_i) t} & (M2) \\ c(x_i, t) = a_1(x_i) + a_2(x_i) t + a_3(x_i) e^{-a_4(x_i) t} & (M3) \end{cases} \quad (4.11)$$

où les coefficients $a_1(x_i)$, $a_2(x_i)$, $a_3(x_i)$ et $a_4(x_i)$ varient avec la position spatiale x_i dans l'image.

La figure 4.6 présente quelques signaux temporels de Rab6-GFP (microscopie WF) correspondant à six positions dans une même séquence non déconvoluée et restaurée par la méthode décrite dans (Boulanger et al. 2010). Les lieux correspondants sont représentés par des croix rouges (Fig. 4.6 (haut)). Le modèle (M3) tracé en rouge sur chacune des figures (signaux 1 à 6) est capable de décrire l'enveloppe convexe inférieure des signaux. Nous remarquons que les coefficients $a_1(x_i)$, $a_2(x_i)$, $a_3(x_i)$ et $a_4(x_i)$ du modèle (M3) sont très différents d'un signal à l'autre et que le passage de vésicules est caractérisé par des pics d'intensité. Nous faisons remarquer que les coefficients des signaux temporels 4 à 6 ne peuvent pas être correctement estimés au sens des moindres carrés. Pour traiter ces situations, il sera préférable de minimiser une norme robuste (Huber 1981) afin de ne prendre en compte qu'un nombre limité d'échantillons temporels. Notons aussi que, après quelques instants, le terme linéaire permet d'expliquer correctement l'évolution du « fond » de l'image (Fig. 4.6). Lorsque la durée de la séquence est trop grande par rapport à la vitesse de décroissance des signaux, il est certain que ce modèle ne convient pas. Cependant, pour les durées d'observation des séquences que nous avons considérées lors de cette étude, le modèle linéaire (M1) d'évolution temporelle nous a permis d'appréhender d'une manière satisfaisante les variations « lentes » du signal pour lesquelles le photo-blanchiment n'est pas seul responsable.

M-Estimateur asymétrique et régression paramétrique.

Dans (Boulanger, Kervrann & Bouthemy 2006, Boulanger, Kervrann & Bouthemy 2009), nous avons considéré le modèle (M1). Les paramètres $\mathbf{a}(x_i) = (a_1(x_i), a_2(x_i))^T$ sont estimés à chaque position x_i à l'aide d'une procédure des « moindres carrés pondérés itérés » qui permet de minimiser le critère suivant :

$$E(\mathbf{a}(x_i)) = \sum_{t=1}^T \rho(f(x_i, t) - a_1(x_i) - a_2(x_i) t) \quad (4.12)$$

4.3 Correction du photo-blanchiment et soustraction de « fond »

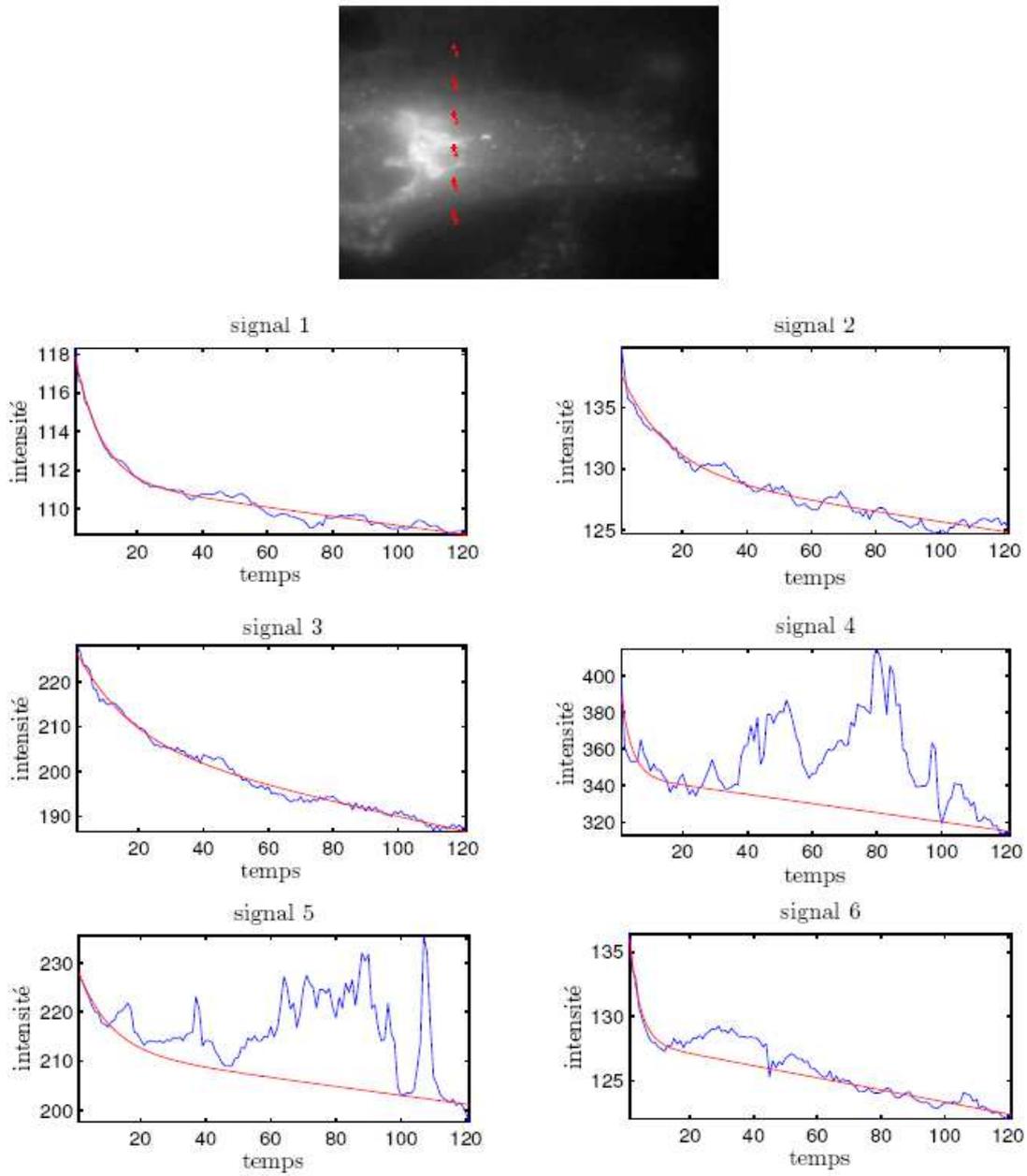


FIG. 4.6. Six signaux temporels (Rab6-GFP) extraits d'une séquence d'images de vidéo-microscopie (WF) réelle restaurée. Chaque signal original est tracé en bleu. Un modèle linéaire avec un terme exponentiel est ajusté manuellement et tracé en rouge.

où T désigne le nombre d'échantillons temporels et $\rho : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction robuste, strictement croissante sur \mathbb{R}_+ et telle que $\lim_{r \rightarrow +\infty} \rho'(r) < +\infty$. Une approximation non-asymptotique de la matrice de variance-covariance $\widehat{\mathbf{C}}(x_i)$ de l'estimateur est également calculée en tout point (Ieng, Tarel & Charbonnier 2004) :

$$\widehat{\mathbf{C}}(x_i) = \frac{\sum_{t=1}^T z(x_i, t) r^2(x_i, t) \sum_{t=1}^T z^2(x_i, t)}{(\sum_{t=1}^T z(x_i, t))^2} \times \left(\begin{array}{cc} \sum_{t=1}^T z(x_i, t) & \sum_{t=1}^T z(x_i, t) t \\ \sum_{t=1}^T z(x_i, t) t & \sum_{t=1}^T z(x_i, t) t^2 \end{array} \right)^{-1} \quad (4.13)$$

où $r(x_i, t) = f(x_i, t) - a_1(x_i) - a_2(x_i) t$ et $z(x_i, t) = \rho'(r(x_i, t))/r(x_i, t)$.

Dans notre situation, le bruit total résulte de la somme des deux composantes $\xi(x_i, t)$ et $v(x_i, t) \geq 0$ correspondant aux vésicules en mouvement. La densité de probabilité du bruit présente donc une forte asymétrie. Pour prendre en compte ce type d'*a priori*, nous avons préconisé une fonction d'influence robuste asymétrique (Ruckstuhl, Jacobson, Field & Dodd 2001, Allende, Frery, Galbiati & Pizarro 2006) :

$$\rho(r) = \begin{cases} 1 - e^{-\frac{r^2}{\lambda\sigma_1^2}} & \text{si } r \leq 0, \\ 1 - e^{-\frac{r^2}{\lambda\sigma_2^2}} & \text{sinon.} \end{cases} \quad (4.14)$$

où les deux paramètres d'échelle notés σ_1 et σ_2 , contrôlent l'asymétrie et λ est une constante positive (fonction de Leclerc (Huber 1981)). Le paramètre d'échelle σ_1 , estimé au sens des moindres carrés à partir des résidus $r(x_i, t)$, est lié à l'amplitude de l'intensité des vésicules. Le second paramètre d'échelle σ_2 correspond au niveau de bruit $n(x_i, t)$ et est généralement estimé à partir des « pseudo-résidus » $(f(x_i, t+1) - f(x_i, t))/\sqrt{2}$ à l'aide d'une procédure d'estimation robuste (« moindres médians »). Finalement, dans les régions de l'image où aucune vésicule n'est présente, nous vérifions $\sigma_1 \approx \sigma_2$.

Modélisation de la cohérence spatiale

Nous avons désiré aussi prendre en compte la cohérence spatiale des paramètres estimés à chaque position spatiale. La régularité est inconnue dans ce cas précis et varie dans l'espace. Comme discuté dans le chapitre 3, nous avons mis en place un schéma d'estimation qui analyse le biais et la variance de l'estimateur $\widehat{\mathbf{a}}(x_i) = (\widehat{a}_1(x_i), \widehat{a}_2(x_i))^T$ en tout point de l'image (Kervrann & Boulanger 2006a, Kervrann & Boulanger 2008).

Au lieu de n'utiliser qu'un seul signal temporel pour estimer les coefficients, nous utilisons un ensemble de signaux temporels 1D collectés dans un voisinage spatial du point x_i . Les signaux ainsi pris en compte par cette procédure forment un tube spatio-temporel noté $\phi_k(x_i)$ et paramétré par son rayon $r_k(x_i)$, $k \in \{1, \dots, K\}$:

$$\phi_k(x_i) = \{x_j : \|x_i - x_j\|_p < r_k(x_i) ; t = 1, \dots, T\} \quad (4.15)$$

où $|\phi_k(x_i)| = T \times (2r_k(x_i) + 1)^d$ le nombre d'échantillons dans le tube si on choisit $p = 1$. En considérant une suite croissante de rayons $(r_k(x_i))_{1, \dots, K}$, nous pouvons construire une collection de tubes emboîtés : $\phi_{k-1}(x_i) \subset \phi_k(x_i)$, $k \in \{1, \dots, K\}$.

Afin de sélectionner un rayon $r_k(x_i)$ au point x_i , nous avons cherché à minimiser le risque local

quadratique $\mathbb{E}[\|\mathbf{a}(x_i) - \widehat{\mathbf{a}}(x_i)\|^2]$, ce qui revient à équilibrer le biais et la variance de l'estimateur $\widehat{\mathbf{a}}(x_i)$ (chapitre 3). Soit $\widehat{\mathbf{a}}_k(x_i)$ le vecteur de paramètres estimé selon la procédure décrite précédemment (Section 4.3.1) en considérant le tube $\phi_k(x_i)$. L'estimation de $\widehat{\mathbf{a}}_k(x_i)$ est initialisée avec $\widehat{\mathbf{a}}_{k-1}(x_i)$ obtenu pour le tube $\phi_{k-1}(x_i)$ de rayon inférieur. Une version vectorielle de la règle d'arrêt de Lepski décrite dans le chapitre 3 et introduite par Polzehl & Spokoiny (1999) permet d'établir une règle de sélection du rayon $r_k(x_i)$. Le tube spatio-temporel qui permet d'équilibrer de manière optimale le biais et la variance de l'estimateur est de la forme :

$$\phi_{k^*}(x_i) = \sup_{k \in \{1, \dots, K\}} \left\{ (\widehat{\mathbf{a}}_k(x_i) - \widehat{\mathbf{a}}_{k'}(x_i))^T \widehat{\mathbf{C}}_{k'}^{-1}(x_i) (\widehat{\mathbf{a}}_k(x_i) - \widehat{\mathbf{a}}_{k'}(x_i)) < \frac{2\eta|\phi_k(x_i)|}{|\phi_k(x_i)|-1}, \right. \\ \left. \text{pour tout } 1 \leq k' < k \right\}.$$

La matrice de variance-covariance étant estimée, le seuil η est alors établi sur la base d'un test de Fisher (Saporta 1978). On notera la différence entre le test construit ici et le test proposé par Polzehl & Spokoiny (1999) qui néglige la corrélation des coefficients $\widehat{a}_1(x_i)$ et $\widehat{a}_2(x_i)$. On peut également remarquer qu'il est impossible d'exprimer cette règle sous la forme d'intervalles de confiance. Ces derniers sont décrits par des ellipsoïdes dont les intersections ne peuvent être calculées analytiquement. L'historique des estimations $(\widehat{\mathbf{a}}_k(x_i))_{k=1, \dots, K}$, est donc mémorisé. Un résultat obtenu avec cette méthode décrite dans (Boulanger et al. 2009), est présenté sur la figure 4.7.

4.3.2 Approche temporelle et non-paramétrique

Dans (Chessel, Cinquin, Bardin, Salamero & Kervrann 2009), nous avons adopté un point de vue différent et nous avons cherché à estimer de manière non paramétrique la composante « fond » en considérant les enveloppes convexes inférieures des signaux temporels. On suppose alors que les variations du signal de type « pics » ou « bosses » doivent appartenir à la composante « objet ». L'enveloppe convexe inférieure est extraite selon le principe des « α -shapes » (Edelsbrunner & Mucke 1994) introduits en géométrie algorithmique (de Berg, van Kreveld, Overmars & Schwarzkopf 1997) et étudiés dans notre cas en dimension 1.

L'idée principale repose sur un sous-échantillonnage adaptatif du signal original 1D $\varphi_0 = f(x_i)$ extrait de la séquence temporelle où $x_i \in \mathcal{I}, i = 1, \dots, N$. Les points pour lesquels la courbe de concavité est inférieure à $\sqrt{\alpha}$ appartiennent à la composante « fond ». On obtient l'enveloppe convexe du signal si $\alpha \rightarrow \infty$. Soit $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ un signal continu et $\mathcal{P} = \{(t, \varphi(t))\}_{t=1, \dots, T}$ l'ensemble des T points $(t, \varphi(t)) \in \mathbb{R}^2$ en échantillonnant φ aux coordonnées entières (instants temporels). Nous décrivons cet ensemble de points dans \mathbb{R}^2 selon le formalisme proposé en géométrie algorithmique.

Définition 1 \mathcal{P}_α^- (respectivement \mathcal{P}_α^+) est un sous-ensemble de \mathcal{P} pour lequel les points de concavité inférieure (respectivement supérieure) à $\sqrt{\alpha}$ ont été supprimés.

En interpolant ces ensembles de points, nous pouvons reconstruire des fonctions continues qui sont des versions simplifiées du signal 1D obtenues à différentes « échelles ».

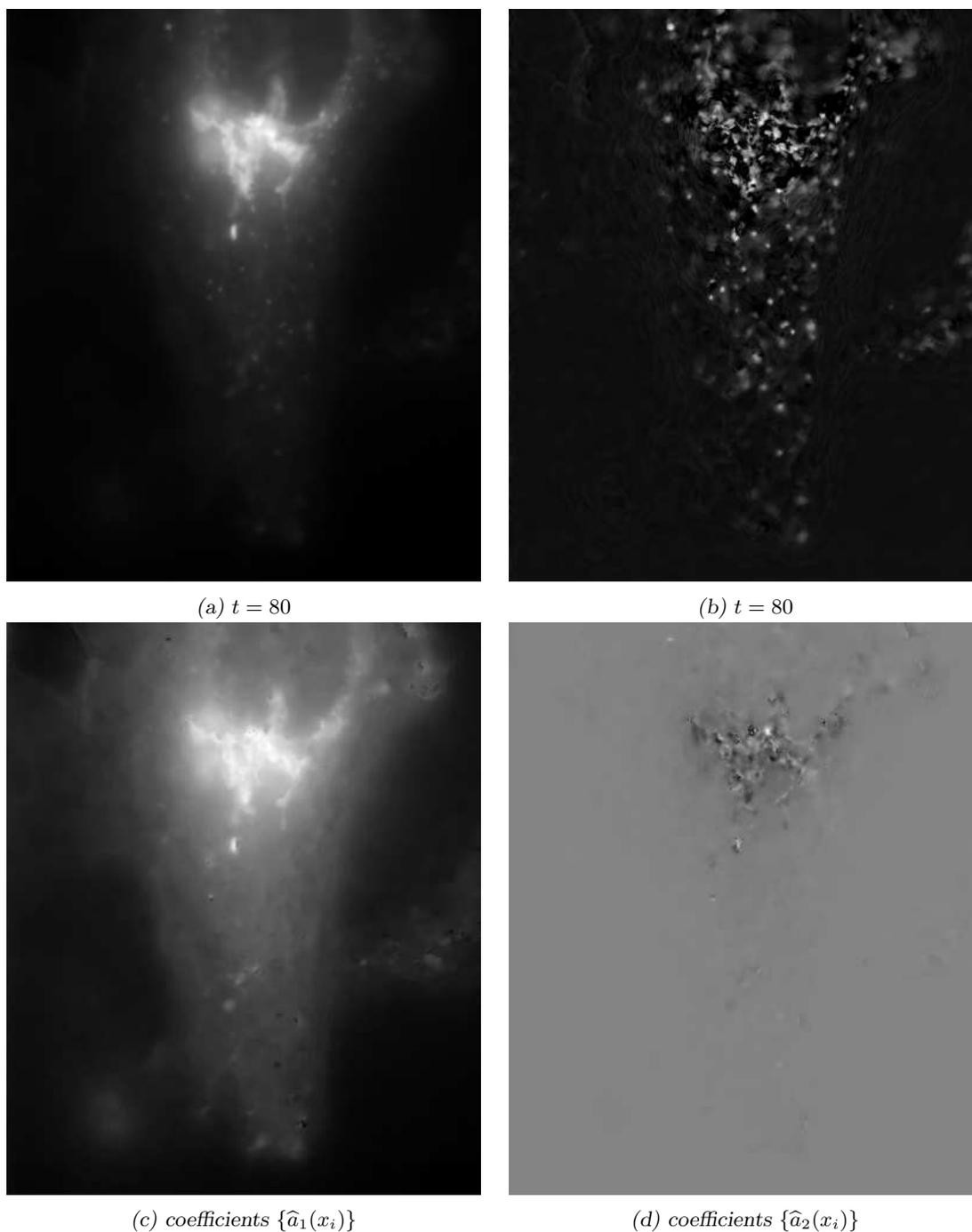


FIG. 4.7. Résultats de l'estimation du fond sur une séquence 2D+T réelle de microscopie WF. (a) image originale ; (b) résidus après soustraction du modèle de « fond » ; (c-d) estimations des coefficients $\{\hat{a}_1(x_i)\}$ et $\{\hat{a}_2(x_i)\}$.

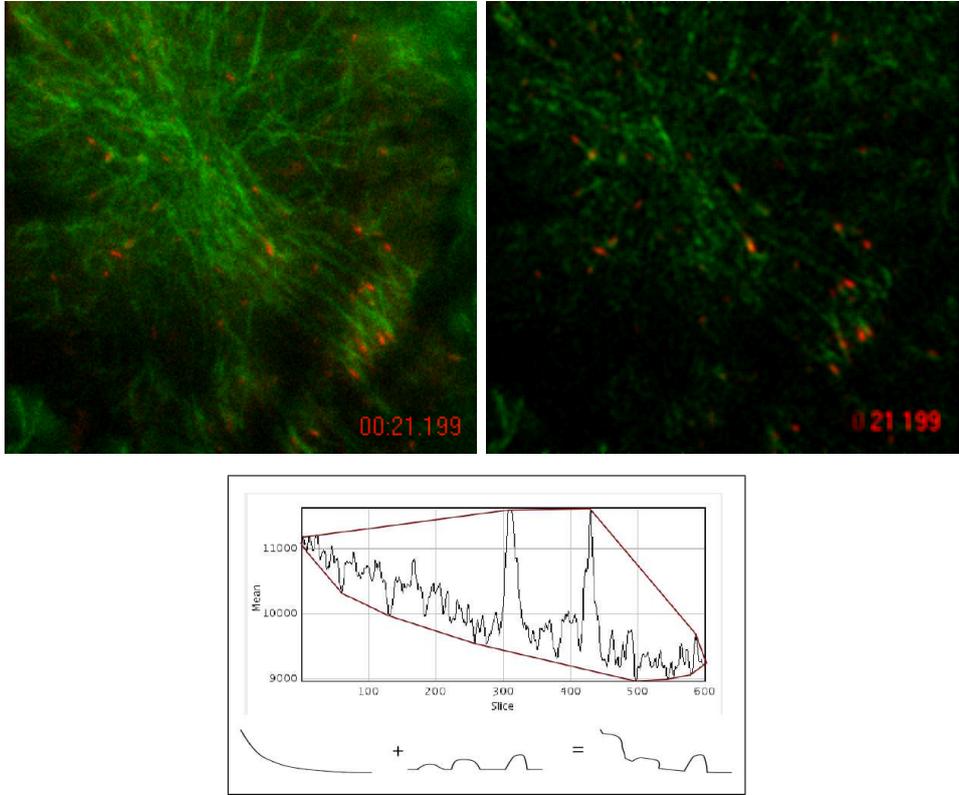


FIG. 4.8. Tub-GFP CLIP-170-mcherry (TIRF) (A. Dimitrov, F. Pérez) : co-localisation de la protéine CLIP170 (rouge) avec les extrémités '+' des microtubules (vert) en séparant les composantes « rapides » et « lentes » du réseau de microtubules (algorithme 'HullkGround').

Définition 2 Soit \mathbb{I} un opérateur d'interpolation vérifiant le principe de maximalité. À l'échelle α , on associe les enveloppes convexes supérieures $\varphi_{\alpha}^{+} = \mathbb{I}(\mathcal{P}_{\alpha}^{+})$ et inférieures $\varphi_{\alpha}^{-} = \mathbb{I}(\mathcal{P}_{\alpha}^{-})$ telles que $\varphi_{\alpha}^{-} \leq \varphi \leq \varphi_{\alpha}^{+}$. Une version continue et simplifiée du signal à l'échelle α est alors définie comme $\varphi_{\alpha} = (\varphi_{\alpha}^{-} + \varphi_{\alpha}^{+})/2$.

L'espace-échelle ainsi défini est invariant aux rotations, translations, inversions de contraste ou décalages de niveaux de gris, mais pas aux changements de contraste d'une manière générale.

En pratique, nous avons utilisé la bibliothèque CGAL (CGAL 2006) conçue en géométrie algorithmique et nous avons considéré une simple interpolation linéaire pour traiter les signaux temporels 1D. L'extraction des composantes pseudo-convexes est effectuée indépendamment en chaque pixel, et le résultat est, de manière étonnante, spatialement régulier sur des images de vidéo-microscopie sans qu'aucune régularité ne soit imposée lors du calcul de la composante convexe. La différence entre la séquence d'images originale et les composantes convexes extraites pour chaque pixel, permet d'identifier la composante « objet » de la séquence. Quelques résultats

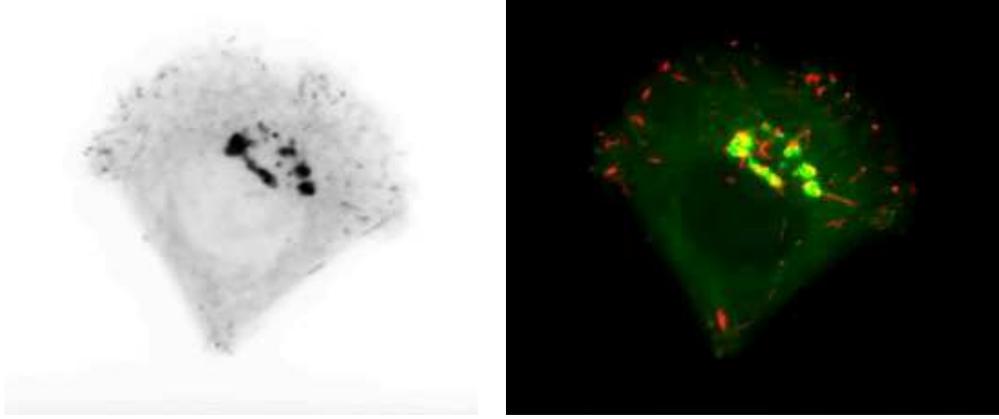


Fig. 4.9. Rab6-GFP (MSDC) : séparation des composantes cytosolique (vert) et vésiculaire (rouge) sur cellule HeLa contrainte par un micro-patron de forme « arbalète » (algorithme 'HullkGround').

de l'algorithme 'HullkGround' (plug-in ImageJ, Java) sont enfin présentés sur les figures 4.8 et 4.9. L'algorithme nécessite l'ajustement d'un seul paramètre α et permet de séparer très rapidement les composantes « lentes » et « rapides » associées aux dynamiques du « fond » et des « objets » dans une séquence d'images de vidéo-microscopie.

Cette approche « discrète » de la convexité n'est pas sans rapport avec l'équation aux dérivées partielles suivante et étudiée par Vese (1999) :

$$\begin{cases} \frac{\partial \varphi}{\partial \tau} = \sqrt{1 + |D\varphi|^2} \min(0, D^2\varphi), & (\tau, t) \in (0, \infty) \times \mathbb{R} \\ \varphi(0, t) = \varphi_0(t), & t \in \mathbb{R}, \end{cases} \quad (4.16)$$

qui permet d'exhiber de manière continue, selon un mouvement par courbure, une famille de fonctions de convexité croissante et ordonnée jusqu'à extraire l'enveloppe convexe du signal original $\varphi_0 : \mathbb{R} \rightarrow \mathbb{R}$. L'existence et l'unicité des solutions de viscosité ont été établies pour cette EDP. Deux flots, convergeant respectivement vers les enveloppes convexes supérieures et inférieures sont ainsi obtenus, coïncidant assez bien avec les définitions de v_α . Néanmoins, des efforts sont encore nécessaires pour complètement établir les relations entre ces deux formalismes relativement différents. Sur le plan de l'implémentation, la mise en place des algorithmes est aussi très différente.

4.3.3 Bilan

À l'aide de ces algorithmes de traitement d'images, nous avons notamment étudié les dynamiques des protéines Rab impliquées dans les voies de transport d'endocytose et d'exocytose. La soustraction de « fond » est une étape qui s'est vite avérée nécessaire pour isoler les composantes membranaire et cytosolique permettant une meilleure compréhension du trafic vésiculaire ou co-localisation des composantes « rapides » et « lentes » des molécules (Fig. 4.8). Plusieurs

points de vue méthodologiques ont été adoptés afin de mettre au point des algorithmes fiables, faciles d'emploi voire rapides. Ces algorithmes sont arrivés à maturité désormais et ont été mis en compétition pour une évaluation objective (Chessel, Pécot, Bardin, Kervrann & Salamero 2009). Selon les applications et les requêtes exprimées par les utilisateurs, il est possible de traiter chaque signal temporel indépendamment (Chessel et al. 2009), de considérer des modèles paramétriques et un contexte semi-local (Boulanger et al. 2009), ou d'adopter un point de vue bayésien et un formalisme markovien (Pécot, Chessel, Bardin, Salamero, Bouthemy & Kervrann 2009). Dans le troisième cas de figure, le « pilotage » des algorithmes associés est néanmoins délicat et nécessite des connaissances *a priori* sur les dynamiques à identifier.

4.4 Détection de changements spatio-temporels en vidéo-microscopie

Dans cette section, nous traitons le problème de la détection significative de changements temporels. Le « fond » est préalablement identifié et soustrait à l'aide d'un des algorithmes que nous venons de décrire dans la section précédente. On s'intéresse désormais aux signaux-images $(v_t)_{t=1, \dots, T}$ où $v_t = (v_t(x_i))_{x_i \in \mathcal{I}}$, correspond au modèle (4.10) de la Section 4.3.1.

4.4.1 Sélection de modèles unidimensionnels

Nous présentons une première contribution (Boulanger et al. 2006, Boulanger et al. 2009) qui s'inspire notamment des travaux sur la « sélection de modèles », décrits dans (Huet 2002, Huet 2006). Il s'agit de déterminer le nombre de composantes nulles dans un vecteur gaussien en minimisant un « critère de vraisemblance pénalisée ». Dans le cas présent, les composantes non nulles coïncident avec les intermédiaires de transport ou spots à détecter. Nous avons adapté cette approche aux signaux temporels, obtenus après soustraction du « fond », en considérant le modèle de signal suivant à la position $x_i \in \mathcal{I}, i = 1, \dots, N$ où $\mathcal{I} \subset \mathbb{R}^2$ est une grille discrète comme précédemment :

$$\mathbf{v}(x_i) = \boldsymbol{\mu}(x_i) + \mathbf{e}(x_i), \quad \mathbf{e}(x_i) \sim \mathcal{N}(0, \boldsymbol{\Sigma}), \quad (4.17)$$

où $\mathbf{v}(x_i) = (v(x_i, 1), \dots, v(x_i, T))^T$ est un signal représenté par un vecteur de dimension T et $v(x_i, t) \in \mathbb{R}, t = 1, \dots, T$. On s'intéresse ici au cas d'un seul signal de longueur T , la même procédure d'estimation étant adoptée pour chaque position spatiale. Dans notre modélisation, la corrélation temporelle des résidus est négligée et on suppose que les variables aléatoires sont indépendantes. Le vecteur \mathbf{e} représente un bruit blanc gaussien de matrice de variance-covariance $\boldsymbol{\Sigma} = v^2 \mathbf{I}_T$ où \mathbf{I}_T est la matrice identité. L'espérance $\mathbb{E}[\mathbf{v}] = \boldsymbol{\mu} = (\mu_1, \dots, \mu_T)^T$ ainsi que la variance v^2 sont inconnues. Désormais, ℓ composantes de $\boldsymbol{\mu}$ sont non nulles. Le vecteur des résidus \mathbf{v} peut être alors ordonné tel que $|v_j| > |v_{j+1}|, j = 1, \dots, T-1$. On considère finalement la collection $\mathcal{M}_T = \{m_1, \dots, m_T\}$ de modèles suivants :

$$\left\{ \begin{array}{l} m_1 = (v_1 \ 0 \ 0 \ \dots \ 0)^T, \\ m_2 = (v_1 \ v_2 \ v_3 \ \dots \ 0)^T, \\ \vdots \\ m_T = (v_1 \ v_2 \ v_3 \ \dots \ v_T)^T. \end{array} \right.$$

On associe à chaque modèle un risque défini par la mesure de Kullback :

$$\mathcal{R}(m_\ell) = \mathbb{E} \left[\mathcal{K}_{\boldsymbol{\mu}, v^2}(\widehat{\boldsymbol{\mu}}_\ell, \widehat{v}_\ell^2) \right] \quad (4.18)$$

où $\mathcal{K}_{\boldsymbol{\mu}, v^2}(\widehat{\boldsymbol{\mu}}_\ell, \widehat{v}_\ell^2)$ est la distance de Kullback et $\widehat{\boldsymbol{\mu}}_\ell$ et $\widehat{v}_\ell^2 = \frac{1}{T} \sum_{j=\ell+1}^T v_j^2$ désignent respectivement les estimateurs au sens du maximum de vraisemblance de $\boldsymbol{\mu}$ et v_ℓ^2 . Le modèle idéal est choisi comme celui minimisant le risque sur l'ensemble de la collection :

$$m^* = \inf_{m_\ell \in \mathcal{M}_T} \mathcal{R}(m_\ell). \quad (4.19)$$

Ceci revient encore à minimiser le critère de vraisemblance pénalisée suivant (Birgé & Massart 2001) :

$$J(m_\ell) = -\mathcal{L}(\mathbf{v}; \boldsymbol{\mu}_\ell, \widehat{v}_\ell^2) + \text{pen}(\ell) \quad (4.20)$$

où $\mathcal{L}(\mathbf{v}; \boldsymbol{\mu}_\ell, \widehat{v}_\ell^2)$ est la log-vraisemblance du modèle $\widehat{\mathbf{m}}_\ell$ et $\text{pen}(\ell)$ une fonction de pénalisation dépendant de la longueur ℓ du modèle. On montre alors que le risque de l'estimateur obtenu est proche du risque idéal (4.18) si la pénalisation est de la forme suivante (Birgé & Massart 2001) :

$$\text{pen}_{MKR}(\ell) = T \left(c_1 \log \left(\frac{T}{\ell} \right) + c_2 \right) \frac{\ell}{T - \ell} \quad (4.21)$$

Notons que les deux constantes « universelles » c_1 et c_2 ont été calibrées par Huet (2002) et fixées à $c_1 = 2$ et $c_2 = 4$. L'auteur propose par ailleurs, une étude comparative entre différentes pénalisations. Parmi celles-ci, nous retrouvons le critère AIC (Akaike 1974) $\text{pen}_{\text{AIC}}(\ell) = \ell$, le critère SIC (Schwarz 1978a) $\text{pen}_{\text{SIC}}(\ell) = \frac{1}{2}\ell \log(T)$ et le critère AMDL (« *Approximate Minimum Length Description* ») introduit par Rissanen (Rissanen 1978) $\text{pen}_{\text{AMDL}}(\ell) = \frac{3}{2}\ell \log(T)$. L'étude numérique menée par Huet (2002) sur une centaine de modèles montre que le risque du critère faisant appel à la pénalisation $\text{pen}_{MKR}(\ell)$ est uniformément plus petit que les critères associés aux autres pénalisations sur des signaux de dimension $T \in [20, 100]$. Nous retenons donc cette pénalisation, et finalement, le modèle est choisi comme celui minimisant le critère de vraisemblance pénalisée suivant :

$$J(M_\ell) = \frac{T}{2} \log(\widehat{v}_\ell^2) + 2T \left(\log \left(\frac{T}{\ell} \right) + 2 \right) \frac{\ell}{T - \ell}. \quad (4.22)$$

Une étude approfondie du critère (4.22) met en évidence l'existence d'un premier minimum local, souvent plus pertinent sur le plan expérimental, et d'un minimum global.

Une fois que les seuils ont été établis pour tous les signaux temporels, nous avons cherché à estimer un seuil unique pour toute la séquence, selon une démarche analogue (Boulanger et al. 2006, Boulanger et al. 2009). Nous présentons un résultat de cette procédure sur la figure 4.10. Quelques dizaines de secondes seulement sont nécessaires pour traiter une séquence de 300 images (697×521 pixels). L'approche de « seuillage automatique » s'avère efficace dans les circonstances analysées et facile à mettre en œuvre. Néanmoins, elle repose aussi sur des hypothèses statistiques assez contraignantes (e.g. indépendance de variables aléatoires gaussiennes), ne permettant pas toujours d'étudier tout comportement dynamique observé dans des séquences de vidéo-microscopie. Nous avons donc proposé, par la suite, des approches non-paramétriques

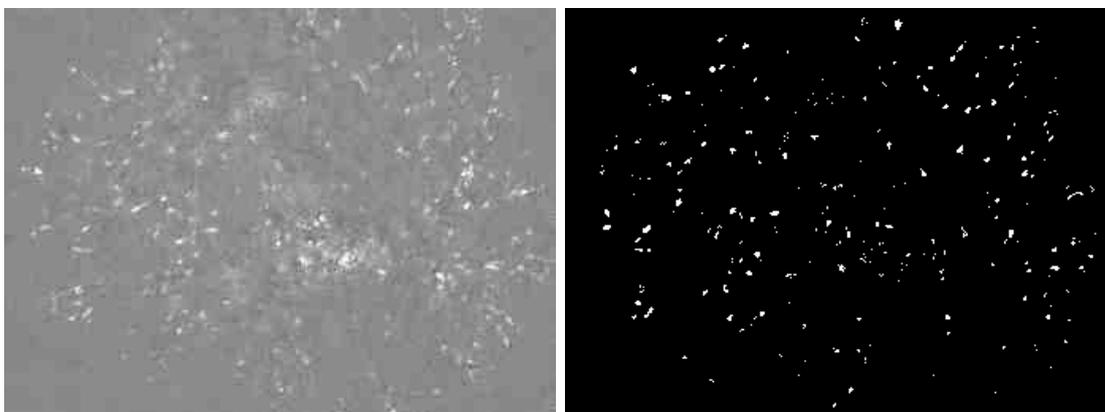


FIG. 4.10. Validation de l'approche par « sélection de modèles » sur une séquence d'images réelles (Rab6-GFP, microscopie WF). L'image v est présentée à gauche à un instant donné $t \in \{1, \dots, T\}$ et la détection correspondante à droite.

permettant de traiter des situations dynamiques plus complexes.

4.4.2 Analyse semi-locale des occurrences dans une paire d'images

Le repérage et la quantification de concentrations de fluorescence transitoires observées en microscope de fluorescence (TIRF, WF) est particulièrement discuté dans cette section. Nous avons ainsi proposé un cadre de détection non-paramétrique de changements temporels, cette fois-ci en analysant des paires d'images consécutives dans une séquence d'images (Boulanger et al. 2009, Chessel, Cinquin, Boulanger, Salamero & Kervrann 2010). Les seuils de détection peuvent dépendre de la position spatiale et sont déterminés sur l'image de référence sans modèle paramétrique de bruit, contrairement à l'approche de « sélection de modèles » précédente. Dans cette étude, il s'agissait d'identifier les spots apparaissant/disparaissant soudainement. Ceux-ci se distinguent des spots animés de mouvements latéraux qui, collectivement, définissent un trafic. Ces événements ponctuels et transitoires sont caractéristiques des mécanismes de recyclage d'endocytose et d'exocytose qui permettent à la cellule de communiquer avec le milieu externe tout au long de son évolution. Nous nous sommes intéressés aux récepteurs Transferrine (TfR) et Langerine dont les expressions sont mesurées au niveau de la membrane plasmique.

La microscopie TIRF permet d'imager la surface de la cellule sur une épaisseur très mince, généralement jusqu'à une profondeur de 100 ou 200 nm (Axelrod 1981). Cette modalité exploite les propriétés de propagation d'ondes évanescentes à l'interface entre deux milieux ayant des indices de réfraction différentes. Cette modalité d'acquisition d'images 2D très rapide (50 ms / image) est très bien adaptée à l'analyse de concentrations transitoires de fluorescence au niveau de la membrane de la cellule. Typiquement, une centaine d'événements (apparition/disparition de spots) par minute sont observés en microscopie TIRF. Néanmoins, il est presque impossible de prédire leurs positions spatiales dans la cellule ou de connaître à l'avance la temporalité de ces événements. Par ailleurs il est question de perturber la cellule afin de faire varier le nombre d'événements transitoires et la temporalité de ces événements.

Des algorithmes automatiques, fiables et rapides de traitement de ces images sont donc nécessaires pour mener de telles études sur un grand nombre de séquences. Nous avons assez vite renoncé à une catégorie d'algorithmes et méthodes de détection de changements temporels (e.g. (Rosin 2002, Radke et al. 2005)), de détection par super-résolution (Thomann, Rines, Sorger & Danuser 2002, Thomann, Dorn, Sorger & Danuser 2003, Berger, Cabal, Fabre, Duong, Buc, Nehrbass, Olivo-Marin, Gadal & Zimmer 2008) et de suivi d'objets (Shi & Thomasi 1994, Sage, Neumann, Hediger, Gasser & Unser 2005, Genovesio et al. 2006, Racine, Hertzog, Jouaneau, Salamero, Kervrann & Sibarita 2006, Racine 2006, Chenouard et al. 2009, Smal, Draegestein, Galjart, Niessen & Meijering 2008, Smal, Meijering, Draegestein, Galjart, Grigoriev, Akhmanova, van Royen, Houtsmuller & Niessen 2008) peu robustes dès que le rapport signal-sur-bruit est faible et l'échantillonnage temporel peu élevé. Les hypothèses de conservation de l'intensité au cours du temps sont rarement valides et un gradient temporel significatif relève tout autant d'un mouvement collectif des spots que du bruit du capteur. Nous avons proposé deux algorithmes basés sur la similarité entre motifs locaux dans une paire d'images pour identifier efficacement ces événements transitoires. L'approche considérée exploite les propriétés de redondance spatiale, temporelle et éventuellement multi-échelle des images, et ne requiert que peu de connaissances *a priori* sur les événements à détecter.

Statistiques des distances extrêmes de similarité

Dans cette étude, nous nous intéressons au Récepteur trans-membranaire Langerine-YFP (Uzan-Gafsou, Bausinger, Proamer, Monier, D., Cazenave, Goud, de la Salle, Hanau & Salamero 2007, Valladeau, Ravel, Dezutter-Dambuyant, Moore, Kleijmeer, Liu, Duvert-Frances, Vincent, Schmitt, Davoust, Caux, Lebecque & Saeland 2000) exprimé de manière stable dans des cellules. Les concentrations transitoires en microscopie TIRF suggèrent que ces événements sont liés à des sites membranaires spécialisés dans l'endocytose ou l'exocytose.

Ce premier algorithme décrit dans (Boulanger et al. 2009) cherche très simplement à associer à un motif d'une image au motif le plus similaire dans l'image suivante dans un voisinage de taille 3×3 ou 5×5 pixels. On se pose alors la question de savoir si cette association est légitime ou si les deux motifs sont mis en correspondance à tort. Pour cela, on considère un motif 2D $\mathbf{v}_t(x_i)$ de taille $\sqrt{n} \times \sqrt{n}$ positionné au site $x_i \in \mathcal{I}$ à l'instant t où \mathcal{I} est le support de l'image et $v_t : \mathcal{I} \rightarrow \mathbb{R}$. Soit un voisinage $V(x_i)$ centré au point x_i dans l'image suivante v_{t+1} à l'instant $t + 1$. À chaque position x_i , nous calculons les distances L_2 entre motifs :

$$\phi_{v_t v_{t+1}}(x_i, x_j) = \|\mathbf{v}_t(x_i) - \mathbf{v}_{t+1}(x_j)\|^2, \quad x_j \in V(x_i) \subset \mathcal{I} \quad (4.23)$$

Soit $\mathcal{D}_t(x_i)$ l'ensemble des distances calculées sur le voisinage $V(x_i)$. En l'absence de changement de contexte, les distances sont des réalisations d'une loi de χ_n^2 avec n degrés de liberté (sous une hypothèse gaussienne). On suppose en effet que $v_t(x_i) \sim \mathcal{N}(0, 1)$ après « stabilisation de variance » (voir Section 4.2). Afin de détecter un changement significatif, nous avons examiné la statistique des distances minimum :

$$\phi_t^{min}(x_i) = \min_{\phi_{v_t v_{t+1}}(x_i, x_j) \in \mathcal{D}_t(x_i)} \phi_{v_t v_{t+1}}(x_i, x_j). \quad (4.24)$$

D'après le théorème de Fisher-Tippett (Embrechts, Kluppelberg & Mikosch 1997), la distance

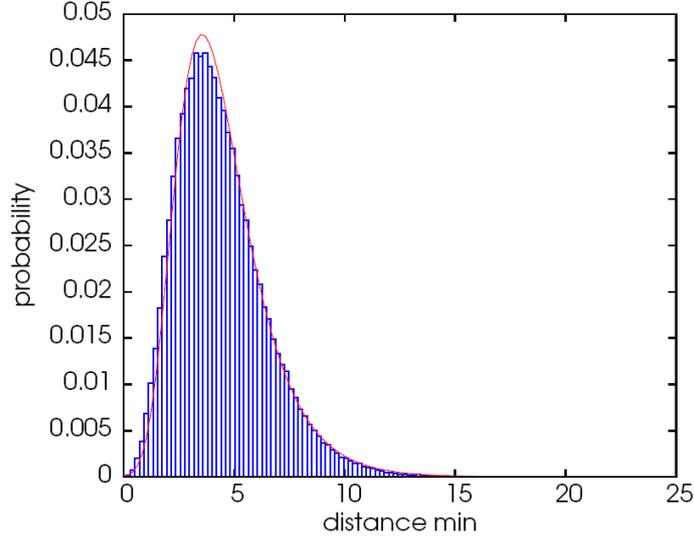


FIG. 4.11. Distribution de la distance minimum entre motifs 3×3 correspondant à des voisinages de taille 3×3 pixels.

minimum suit une distribution de valeurs extrêmes généralisées (GEV) (Figure 4.11) :

$$GEV(z, \alpha, \beta, \kappa) = \begin{cases} \exp - \left(1 - \kappa \left(\frac{z - \beta}{\alpha} \right) \right)^{1/\kappa} & \text{si } \kappa \neq 0 \\ \exp - \left(1 - \exp \left(\frac{z - \beta}{\alpha} \right) \right) & \text{si } \kappa = 0 \end{cases} \quad (4.25)$$

Dans nos expériences, les paramètres de la distribution sont estimés par maximum de vraisemblance (Morrison & Smith 2002). Une décision finale est prise à chaque pixel sur la base de tests multiples et repose sur des hypothèses statistiques pas toujours simples à vérifier sur le plan expérimental. Une illustration de cette méthode de détection est présentée sur la figure 4.12.

Détection non-paramétrique multi-échelle en microscopie TIRF

La microscopie TIRF étant une modalité très rapide d'imagerie 2D, nous avons aussi proposé d'adapter l'algorithme 4 de détection de changements temporels décrit dans la Section 3.3 (voir aussi (Chessel et al. 2010)). L'avantage de cet algorithme réside dans sa robustesse au bruit et aux mouvements résiduels. L'approche est complètement non-paramétrique et la taille des spots est prise en compte dans la modélisation pour limiter le nombre de fausses alarmes. Les pixels qui présentent un niveau de fluorescence trop faibles sont écartés car localisés sous la membrane. Une première illustration est présentée sur la figure 4.13.

Dans une première étude de cas, nous nous sommes intéressés à la molécule Langerine dont le trafic est contrôlé par la protéine Rab11. Les séquences d'images ont été acquises à une résolution temporelle de 100ms / image sur un microscope Nikon TE2000 inversé équipé d'un bras TIRF

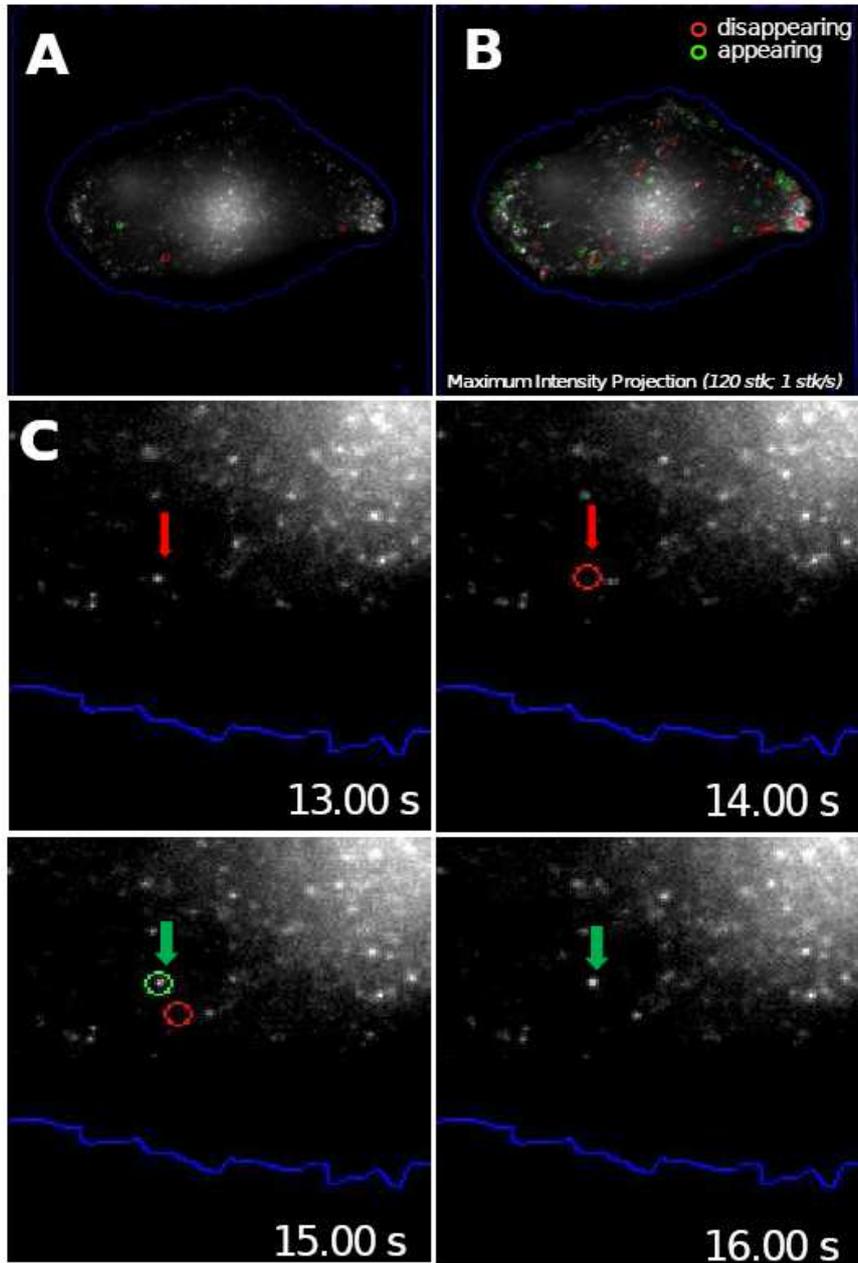


FIG. 4.12. Exemple de validation en microscopie 3D+T (WF, 1 volume 3D / seconde) de l'analyse des statistiques des distances extrêmes de similarité. Détection de changements sur une image 2D (A) affichage des détections sur l'ensemble du volume 3D (projection d'intensité maximum) (B) à un instant donné. Les cercles vert et rouge identifient l'apparition et la disparition soudaines de spots (C).

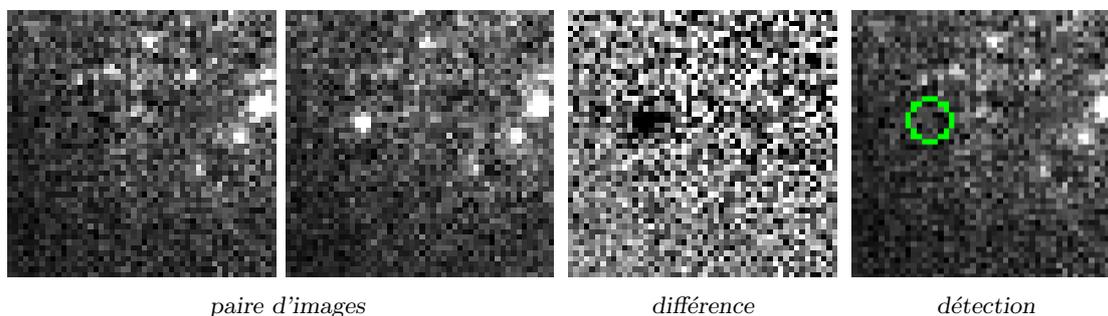


FIG. 4.13. Détection de changements temporels significatifs sur une paire d'images (extrait) de microscopie TIRF (Langerine-YFP).

(ouverture 100x TIRF, objectif (NA=1.49)) sur une durée de 1 minute au total. Sur la figure 4.14 (bas gauche), nous avons reporté l'ensemble des évènements transitoires détectés par l'algorithme pour chacune des molécules observée en « *dual-view* ». L'étude en cours vise désormais à analyser la signalisation observée en chacun des lieux précis repérés automatiquement par notre algorithme (Fig. 4.14 (bas droite)).

Dans une seconde étude de cas, nous avons étudié le récepteur trans-membranaire de recyclage Transferrine (TfR) (Bleil & Bretscher 1982, Grant & Donaldson 2009) dont l'expression et la fonction sont directement liées à la division cellulaire et au métabolisme du fer au niveau cellulaire. Les cellules M10 (« micro-patternées » / forme arbalète) sont transfectées transitoirement avec un plasmide encodant la TfR fusionnée avec la PhLuorin (Grant & Donaldson 2005). La figure 4.15 présente sur une même image (projection selon l'axe temporel) toutes les détections calculées avec cet algorithme et étiquetées par un expert. La précision obtenue est de 85% avec un taux de fausses alarmes inférieur à 10%, ce qui semble suffisamment robuste pour une exploitation plus intensive dans plusieurs études de cas.

4.5 Estimation de trafic par « Tomographie de Réseaux »

La technique de marquage avec la protéine GFP « Green Fluorescent Protein » et la vidéo-microscopie à fluorescence sont des outils d'investigation permettant d'observer des dynamiques de protéines et des interactions moléculaires dans des cellules vivantes. Plus spécifiquement, notre objectif est d'élucider le rôle des protéines de la famille des Rab (petites GTPases) au sein de complexes multi-protéiques (Fig. 4.16). Chaque membre de cette famille existe sous différents états dynamiques dans la cellule : ils échangent continûment entre le cytosol et le compartiment « donneur » ; ils sont associés aux intermédiaires de transport (vésicules) acheminés jusqu'au compartiment donneur en empruntant un réseau de microtubules ; ils s'associent de manière transitoire aux membranes du compartiment « accepteur » (Chavrier & Goud 1999, Seabra, Mules & Hume 2002). Sur le plan expérimental, les travaux menés en biologie cellulaire visent à évaluer le rôle des moteurs moléculaires dans la dynamique des Rab en inactivant certains d'entre eux par interférence ARN, par photo-perturbation (FRAP/4D) en imposant des contraintes sur

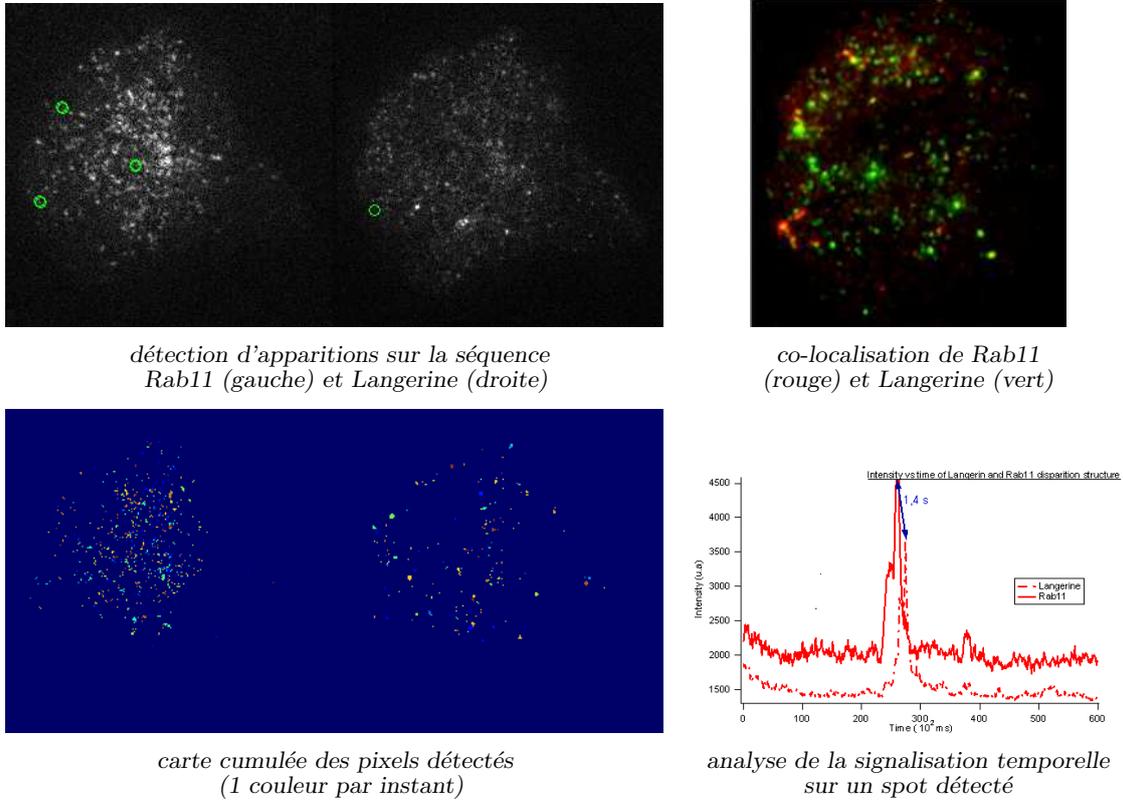


FIG. 4.14. Analyse des évènements transitoires et de la co-localisation des molécules Rab11 et Langerine dans l'espace et dans le temps en microscopie TIRF.

la forme de la cellule à l'aide de « micro-patrons » (They, Racine, Pepin, Piel, Chen, Sibarita & Bornens 2005) ou en perturbant la cellule via des drogues (e.g. nocodazole vs. dépolymérisation des microtubules). Les descripteurs calculés pour chaque condition doivent permettre de mieux préciser les relations spatio-temporelles entre les membranes vésiculaires, les machines responsables du trafic intra-cellulaire, les moteurs moléculaires et les éléments du cytosquelette (Fig. 4.17).

L'analyse du transport membranaire et du trafic vésiculaire pose des questions difficiles sur le plan du traitement de l'information et de l'image. Typiquement, les cellules qui expriment la protéine GFP-Rab6 mettent en évidence des vésicules (intermédiaires de transport) se déplaçant de manière hétérogène le long du réseau de microtubules, propulsées par des moteurs moléculaires (Hirokawa 1998), selon une voie rétrograde, depuis l'appareil de Golgi jusqu'au reticulum endoplasmique (RE) (White, Johannes, Mallar, Girod, Stephan, Reinsh, Keller, Tzschaschel, Echard, Goud & Stelzer 1999, Girod, Storrie, Simpson, Johannes, Goud, Roberts, Lord, Nilson & Pepperkok 1999). Comment peut-on suivre plusieurs dizaines ou centaines d'objets aussi similaires et en interaction d'une manière globale? Chaque vésicule est par ailleurs animée de

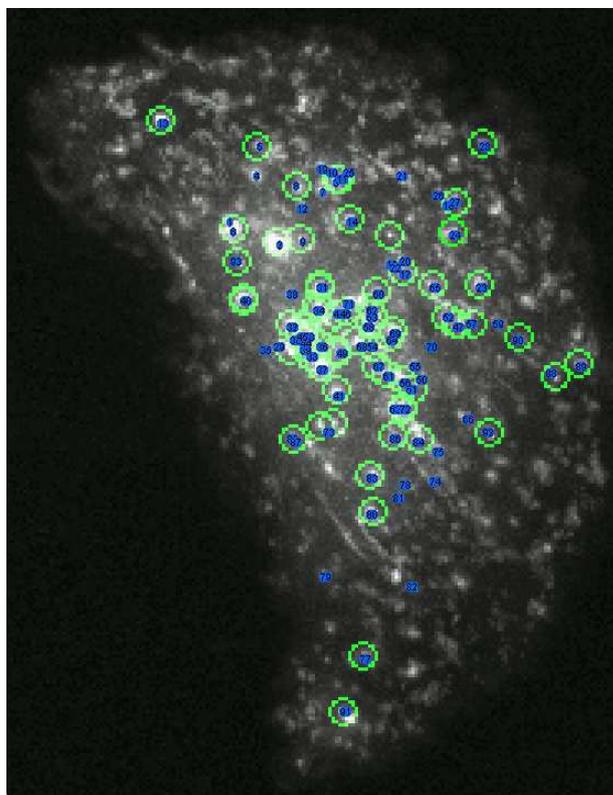


FIG. 4.15. Projection d'une séquence Tfr-PH1 en microscopie TIRF. Les apparitions automatiquement détectées correspondent aux cercles verts. Les événements étiquetés en bleu correspondent aux apparitions identifiées manuellement.

mouvements très complexes : mouvements bi-directionnels principalement dans le sens des extrémités « plus » des microtubules, variations de vitesse importantes entre particules, événements stochastiques en dehors de l'appareil de Golgi... Les mécanismes moléculaires responsables de cette complexité dépendent partiellement de l'identification des moteurs moléculaires (kinesin et dynein) assurant le déplacement des vésicules le long du cytosquelette (Fig. 4.18).

Compte tenu des difficultés des méthodes de filtrage temporel stochastique (trop grand nombre d'objets en forte interaction, modèles dynamiques difficiles à obtenir à partir des connaissances de l'expert, problème d'associations des mesures) (Chenouard et al. 2009) ou de stratégies « connexionnistes » (Bonneau, Dahan & Cohen 2005, Sbalzarini & Koumoutsakos 2005, Racine et al. 2006) ou kymographes (Racine, Saschse, Salamero, Fraasier, Trubuil & Sibarita 2007), nous avons proposé une approche originale pour estimer le trafic sans pour autant recourir au suivi individuel des objets en mouvement (Pécot, Boulanger, Kervrann & Bouthemy 2007, Pécot, Kervrann & Bouthemy 2008b, Pécot, Kervrann & Bouthemy 2008a). Le trafic membranaire est alors assimilé à un trafic routier (Tamin & Willumsen 1989, Bell 1991) ou un trafic sur un réseaux de télécommunications, ce qui permet alors d'exploiter les techniques récentes en « To-

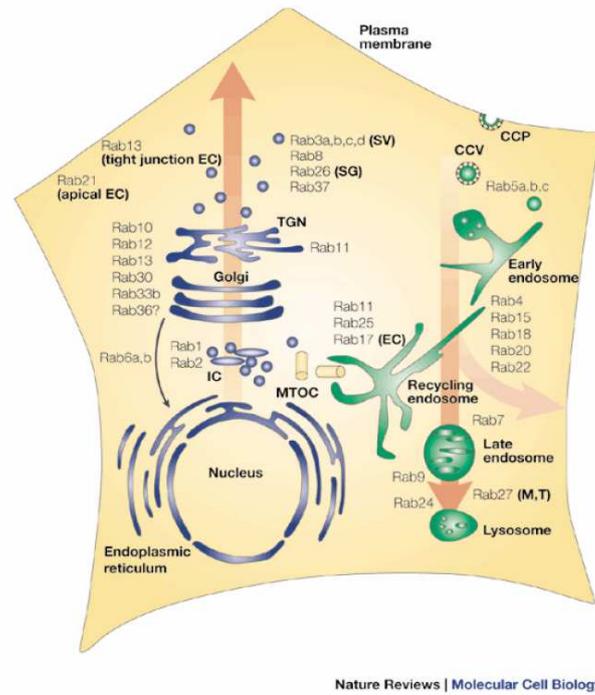


FIG. 4.16. Représentation des domaines des protéines Rabs dans la cellule.

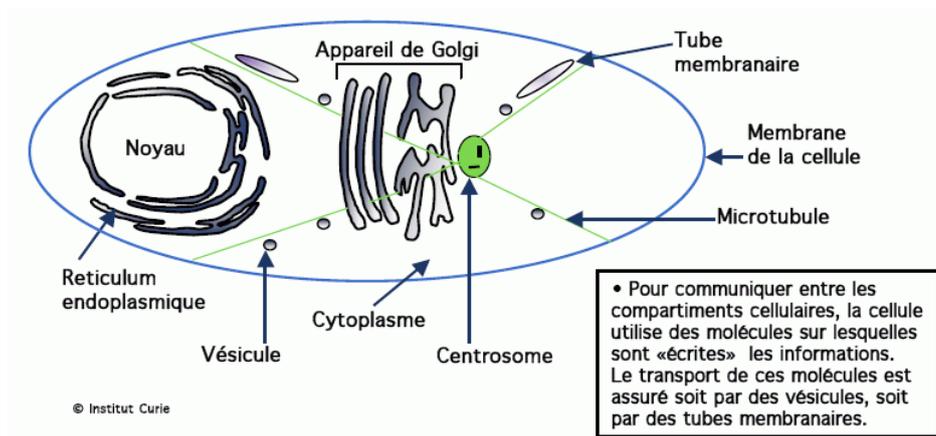


FIG. 4.17. Description élémentaire d'une cellule.

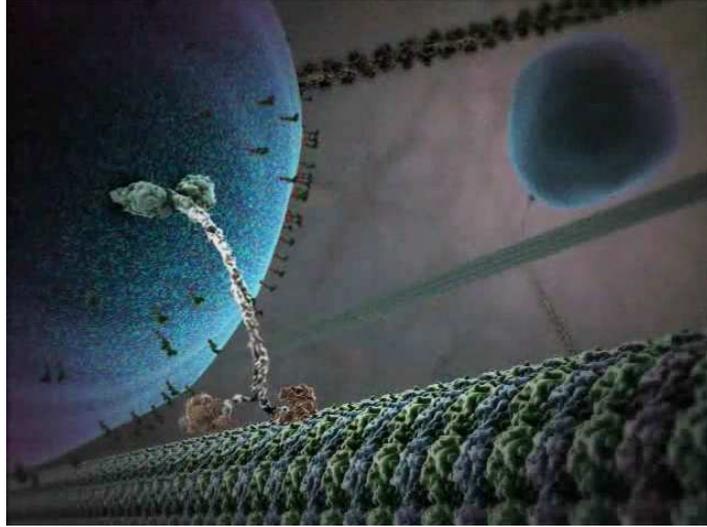


Fig. 4.18. Image extrait d'un film d'animation (droite) illustrant la « marche » d'un moteur moléculaire sur un microtubule et arrimé à une vésicule (« the inner life of the cell », équipe BioVisions, Université de Harvard).

mographie de Réseaux » (Castro, Coates, Liang, Nowak & Yu 2004) introduites par Vardi (1996) puis reprises par Boyd, Meloche & Vardi (1999). L'estimation des trajectoires complètes de tous les objets n'est en effet pas toujours nécessaire à la compréhension et la mesure de l'activité cellulaire. Estimer les flots de données entre régions « origine » et « destination » peut s'avérer plus pertinent en la circonstance.

Dans cette description du problème de suivi, faisant notamment l'objet de la thèse de Thierry Pécot que j'ai encadrée, la scène dynamique est constituée de particules en mouvement se déplaçant sur un réseau de microtubules, décrit par un réseau de régions inter-connectées. Dans ce graphe $\mathcal{G}(\mathcal{E}, \mathcal{V})$ composé de $N_{\mathcal{V}}$ sommets et de ℓ arêtes, où \mathcal{E} désigne l'ensemble des arêtes, et \mathcal{V} l'ensemble des sommets, chaque nœud représente une région biologique sélectionnée par l'expert. Une connexion entre deux nœuds est un chemin et chaque chemin est établi à partir de plusieurs liens (arêtes unidirectionnelles) connectant plusieurs « routeurs » ou nœuds intermédiaires. Il s'agit par la suite d'inférer la nature statistique du réseau à partir de mesures locales obtenues au niveau des nœuds intermédiaires. Ces mesures permettent d'évaluer le nombre de vésicules initialement localisées à un sommet origine et arrivant à un sommet destination en parcourant un chemin, parmi toutes les paires origine-destination (OD) possibles du réseau, c'est-à-dire $c = N_{\mathcal{V}}(N_{\mathcal{V}} - 1)$ paires OD. L'objectif final est d'estimer combien de particules issues d'un nœud « source » atteignent un nœud « destination », en empruntant un chemin quelconque dans le graphe.

Soit $z_{j,t} \in \mathbb{N}$ le nombre de vésicules « transmises » à l'instant t sur la paire origine-destination (OD) $n^{\circ}j$, $j = 1, \dots, c$. Les mesures $\mathbf{y}_t = (y_{1,t}, \dots, y_{\ell,t})^T$ avec $y_{e,t} \in \mathbb{N}$, représentent le nombre de vésicules qui transitent d'un sommet à un autre à l'instant t . Nous modélisons le problème

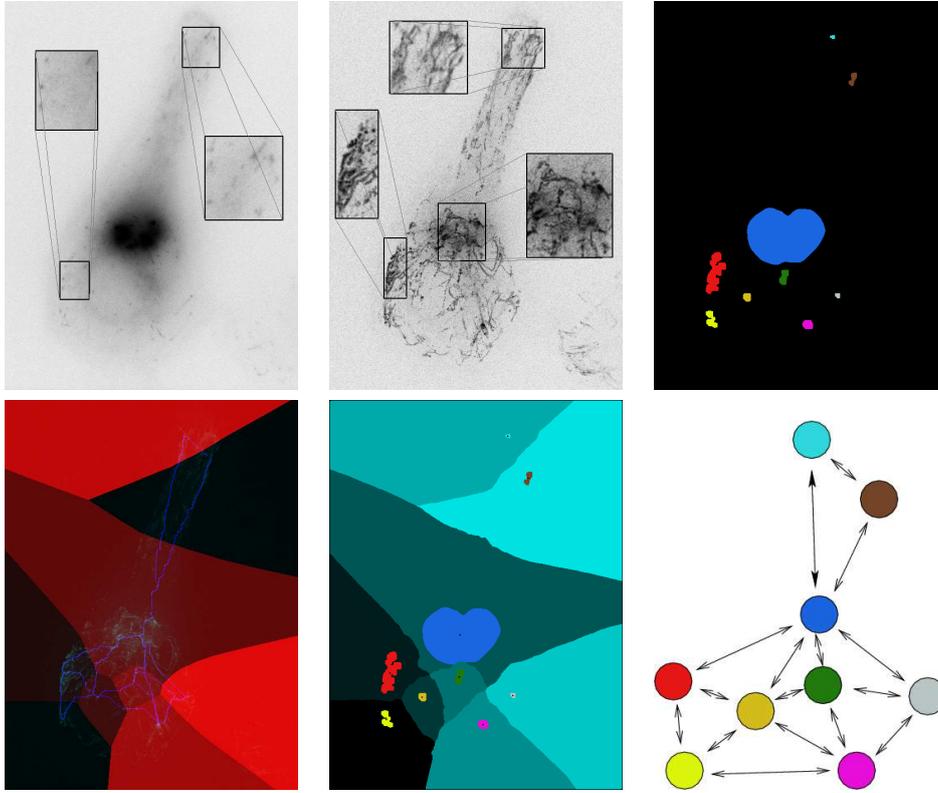


FIG. 4.19. Analyse de trafic de Rab6-GFP. Première ligne : gauche) une image 2D extraite de la séquence de vidéo-microscopie WF ; centre) image Π_v de projection d'intensité maximale selon l'axe temporel ; droite) segmentation des zones d'accumulation de fluorescence. Deuxième ligne : gauche) chemins les plus courts au sens géodésique sur Π_v reliant les régions détectées ; centre) partition géodésique de l'image Π_v ; droite) représentation graphique des origines et destinations.

d'estimation de trafic comme un problème linéaire : $\mathbf{Y} = \mathbf{AZ}$ où $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_T)$ est la matrice des mesures et $\mathbf{z} = (\mathbf{z}_1, \dots, \mathbf{z}_T)$ avec $\mathbf{z}_t = (z_{1,t}, \dots, z_{c,t})^T$, représentent les flux inconnus origine-destination pour une séquence temporelle de longueur T . L'algorithme Dijkstra (1959) est conventionnellement recommandé pour identifier le chemin le plus court dans le graphe, associé à chaque paire OD. La matrice de routage \mathbf{A} de dimension $\ell \times c$ est alors une matrice binaire dont les éléments $a_{ij} = 1$ si l'arête i appartient au chemin de la paire OD $n^o j$, et 0 sinon. Dans le domaine des télécommunications, cette matrice est facilement identifiée puisque le réseau est parfaitement connu. Dans notre problématique en vidéo-microscopie, il s'est avéré nécessaire de réviser la définition de \mathbf{A} , qui est également une inconnue du problème. Dans (Pécot et al. 2008a), plusieurs chemins sont ainsi envisagés pour acheminer les molécules d'une origine à une destination. La probabilité d'emprunter un chemin est alors proportionnelle à la longueur au sens d'une métrique géodésique calculée sur une image qui « résume » l'information dynamique. Nous avons

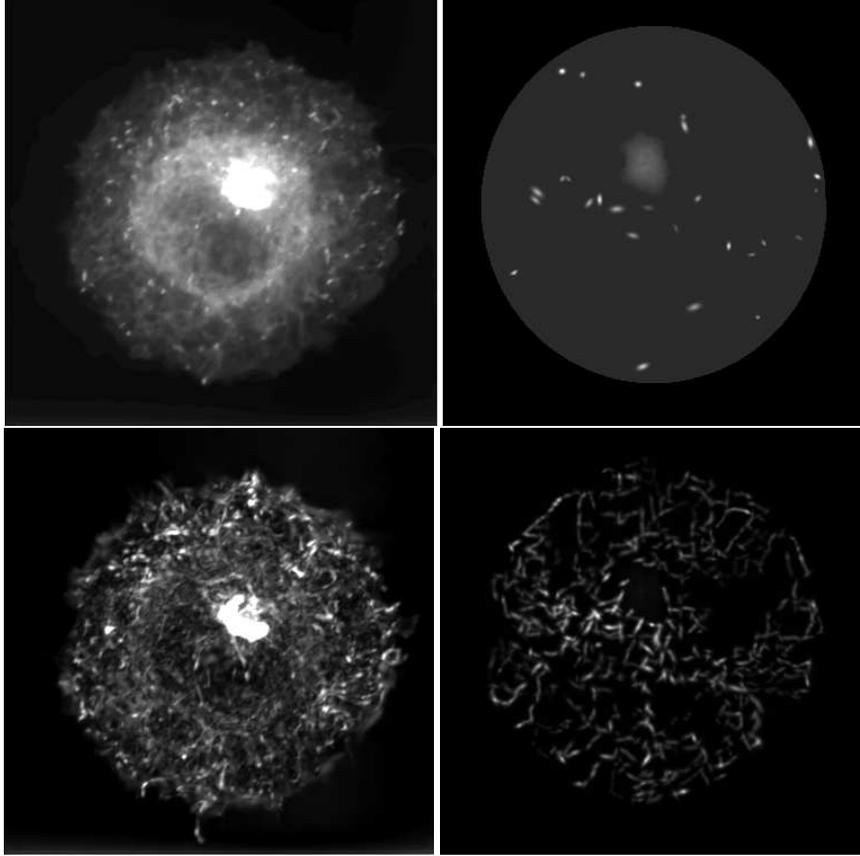


FIG. 4.20. Simulation par « Tomographie de Réseaux » du comportement dynamique de Rab6 sur une cellule de forme ronde (haut) et comparaison des cartes Π_v « résumant » les trajectoires des vésicules.

notamment calculé les chemins les plus courts (Cohen & Kimmel 1997) sur l'image de projection d'intensité maximale Π_v définie telle que : $\Pi_v(x_i) = \sup_{t=1, \dots, T} v_t(x_i)$ où $v_t : \mathcal{I} \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ est une image de la séquence temporelle (obtenue après soustraction du « fond » (voir Section 4.3)) et $x_i \in \mathcal{I}, i = 1, \dots, N$ désigne un pixel de l'image. Une illustration de cette démarche est présentée sur la figure 4.19.

Expérimentalement, la longueur T des séquences est relativement courtes puisqu'elles n'excèdent pas quelques centaines d'images. On se pose donc le problème de l'estimation des paires OD « en moyenne », c'est-à-dire la proportion de vésicules à atteindre une cible connaissant ou non l'origine et en considérant une matrice \mathbf{A}_{Π_v} non binaire. En définitive, le problème de l'estimation des flux origine-destination revient à résoudre un problème d'optimisation pour lequel c est plus grand que ℓ :

$$\min_{\bar{z}} \|\bar{y} - \mathbf{A}_{\Pi_v} \bar{z}\|^2 \quad \text{sous la contrainte} \quad \bar{z}_j \geq 0, \forall j \in \{1, \dots, c\} \quad (4.26)$$

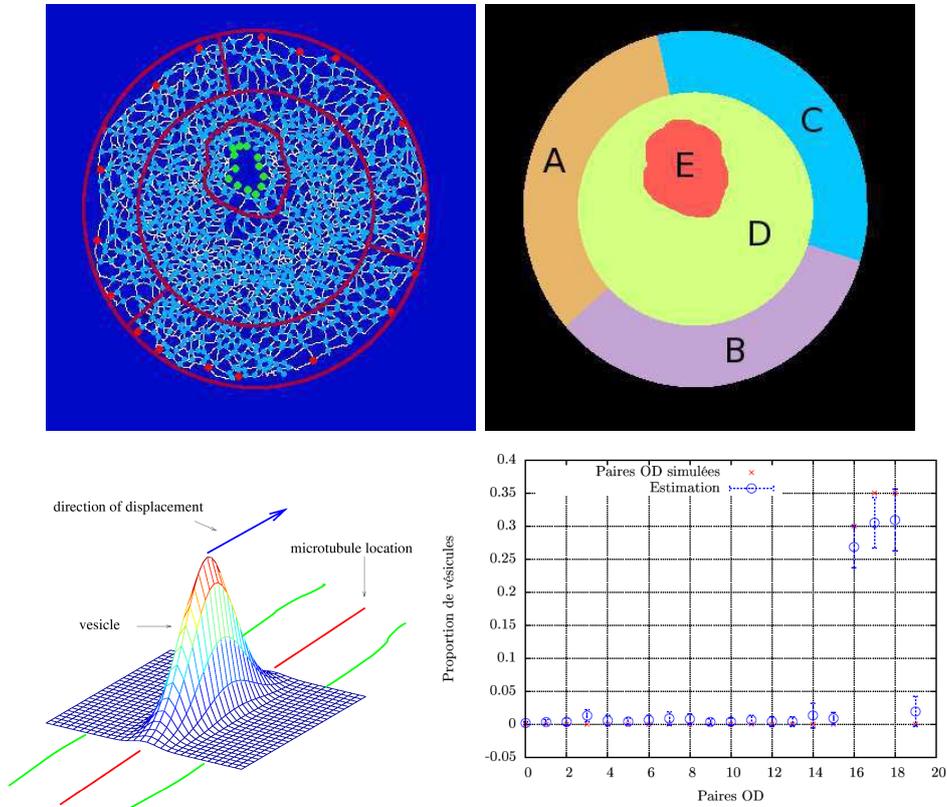


FIG. 4.21. Estimation par « Tomographie de Réseaux » des 20 paires origines-destination (deuxième ligne : droite) correspondant à une segmentation en 5 régions (première ligne : gauche) sur une séquence simulée (voir Fig. 4.20). Les origines et les destinations correspondent aux labels 'vert' et 'rouge' superposés sur l'image de gauche (première ligne) qui illustrent un réseau réel de microtubules. Les vésicules sont assimilées à des spots gaussiens anisotropes (deuxième ligne : gauche).

où $\bar{z} = (\bar{z}_1, \dots, \bar{z}_\ell)^T$ correspondent aux proportions inconnues et $\bar{y} = (\bar{y}_1, \dots, \bar{y}_\ell)^T$ aux moyennes temporelles calculées au niveau des arêtes du graphe (Pécot et al. 2008a) :

$$\bar{y}_e = \frac{1}{T} \sum_{t=1}^T y_{e,t}, \quad \forall e \in \{1, \dots, \ell\}. \quad (4.27)$$

Ces moyennes temporelles sont établies à partir des échanges de fluorescence entre des régions de l'image (e.g. cellules de Voronoï) qui partitionnent le domaine de l'image (Fig. 4.19). Des contraintes supplémentaires sont nécessaires pour résoudre ce problème. Vardi (1996) préconisa des contraintes statistiques pour caractériser le trafic (loi de Poisson). L'hypothèse Poisson n'étant pas très réaliste dans notre cas, nous avons imposé une contrainte de parcimonie

puisque le trafic est observé sur un faible nombre de paires OD. En introduisant la norme $\|\bar{z}\|_0 = \sum_{i=1}^c \mathbb{1}[\bar{z}_i \geq 0]$ puis la norme $\|\bar{z}\|_1 = \sum_{i=1}^c |\bar{z}_i|$, le problème d'optimisation convexe à résoudre est de la forme :

$$\min_{\bar{z}} \|\bar{y} - \mathbf{A}_{\Pi_o} \bar{z}\|^2 + \lambda_{OD} \|\bar{z}\|_1 \quad \text{sous la contrainte} \quad \bar{z}_j \geq 0, \forall j \in \{1, \dots, c\} \quad (4.28)$$

où λ_{OD} est une constante positive. Dans notre étude, nous pouvons aussi imposer les origines et les destinations préalablement localisées (Pécot, Kervrann, Bardin, Goud & Salamero 2009) et rechercher seulement les proportions de trafic. Le problème d'optimisation est bien plus simple car le nombre de paires OD est limité. Avec cette approche, nous bénéficions du fait qu'il n'est pas utile de suivre chaque vésicule qui transite d'un nœud à l'autre. Seul le nombre de particules à un nœud du graphe est important. Un exemple d'estimation des paires origine-destination est présenté sur la figure 4.21 pour une séquence simulée présentée sur la figure 4.20. L'analyse sur des cas réels significatifs est en cours d'étude actuellement.

Par ailleurs, nous avons également montré dans (Boulangier et al. 2009) que ce modèle dynamique permet de réaliser la synthèse d'images selon une démarche qui se distingue des méthodes probabilistes du type Monte Carlo ou de méthodes stochastiques décrites par Wang & Zhu (2002) voire relatives à une modélisation des images par des ensembles aléatoires (Matheron 1975, Gousseau 2000). Dans le cadre de la thèse de Thierry Pécot, ces images simulées prennent désormais en compte des contraintes bio-physiques et permettent ainsi de générer des séquences temporelles pour lesquelles le mouvement des objets est assez réaliste pour ce type de trafic (Fig. 4.20). Elles s'avèrent également pertinentes pour évaluer les algorithmes de suivi plus conventionnels si besoin.

4.6 Conclusion

Sur le thème de l'analyse de dynamiques moléculaires observées en vidéo-microscopie de fluorescence, nous avons proposé plusieurs contributions qui visaient à estimer le photo-blanchiment, les composantes « lentes » et « rapides » et parfois des changements spatio-temporels significatifs révélateurs d'un trafic (protéines de la famille des Rabs) ou d'évènements transitoires (récepteurs Langerine et transferrine). Plusieurs cadres mathématiques ont ainsi été abordés : sélection de modèles, estimation non-paramétrique, géométrie algorithmique, équations aux dérivées partielles, statistiques des valeurs extrêmes. Selon les cas, nous nous sommes intéressés aux signaux temporels ou à des contextes spatio-temporels. Nous avons eu le souci de proposer des algorithmes rapides, fiables et simples d'utilisation par des biologistes.

Pour conclure ce chapitre, je ferai enfin mention de l'approche bayésienne (Champs Aléatoires Conditionnels) développée dans le cadre de la thèse de Thierry Pécot, qui permet d'estimer les composantes membranaire (« objets ») et cytosolique (« fond ») d'une part et, d'autre part, de détecter les vésicules en mouvement (Pécot et al. 2009). Les champs markoviens modélisent l'information contextuelle, très locale et sont bien adaptés en vision par ordinateur. La modélisation directe de la distribution *a posteriori* dans le cadre des champs aléatoires conditionnels (Lafferty et al. 2001, Kumar & Hebert 2006) permet une prise en compte du caractère non local des interactions, et permet une certaine flexibilité pour mêler variables observables et variables

cachées. La modélisation markovienne proposée est relativement flexible et offre des perspectives intéressantes. Ce type d'approche est par ailleurs recommandé pour détecter les objets en mouvement dans une vidéo d'une manière générale (Criminisi, Cross, Blake & Kolmogorov 2006, Sun, Zhang, Tang & Shum 2006). Néanmoins, la minimisation d'une telle énergie peut s'avérer délicate et l'approche nécessite l'ajustement de paramètres, une problématique qu'il s'agira d'aborder à l'avenir si ce type de modélisation devenait indispensable pour analyser le trafic.

Quelques perspectives

Ces dernières années, nous avons opté pour un cadre de modélisation non-paramétrique afin de répondre à de nombreuses questions en analyse d'images, mais également à des questions posées par nos collaborateurs biologistes qui exploitent la microscopie. Les réflexions sur la modélisation d'images que je viens d'exposer dans ce document ont donc été menées en partenariat avec des laboratoires de biologie fondamentale pour lesquels la microscopie est un instrument technologique d'investigation. L'image permet d'analyser des mécanismes spatiaux et temporels qu'il s'agit de décrypter, à l'échelle d'une cellule vivante. Par conséquent, il nous a semblé plus prudent d'analyser les informations sans « trop » d'*a priori*, afin de leur permettre d'avancer des thèses plus objectives en biologie. Cet investissement devrait se poursuivre dans les années à venir et nous serons amenés à traiter des questions assez variées. Le problème de l'imagerie « super-résolution » sera par exemple abordé toujours en relation avec les avancées technologiques en biophotonique.

Parallèlement, j'ai eu l'opportunité d'aborder la tomographie électronique, davantage utilisée pour analyser la structure des éléments de la cellule avec une résolution de 5-10 nanomètres. L'obtention de volumes 3D par cryo-tomographie électronique (cryo-ET) et leur traitement numérique nécessitent des moyens très conséquents en termes de calculs et de visualisation. Il s'agit de reconstruire des structures de la cellule en 3D, à partir de vues 2D inclinées d'échantillons congelés-hydratés. Les volumes obtenus présentent un très faible rapport signal-sur-bruit et les méthodes manuelles de segmentation couramment employées se révèlent inefficaces. Les acquisitions originales sont obtenues sous forme de « piles » de plusieurs dizaines d'images de 2024×2024 ou 4096×4096 pixels (codage sur 32 bits) (résolution de quelques nm), ce qui représente plusieurs gigabits pour un simple jeu de données. Le calcul des volumes à l'aide d'algorithmes dédiés de segmentation par modèles de Markov (« *graph-cuts* » actuellement à l'étude) nécessite des développements particuliers afin de pouvoir analyser ces images de grande taille. Dans le cadre de la thèse de Sophie Blestel, ces algorithmes sont appliqués depuis quelques mois à l'analyse

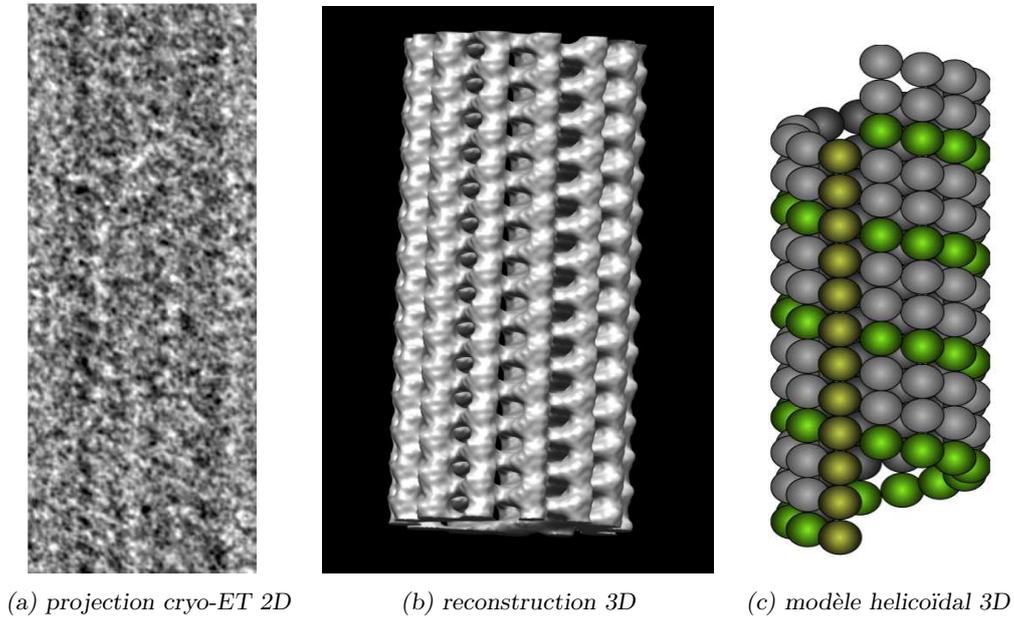


Fig. 5.1. Reconstruction 3D d'un microtubule à partir d'une projection cryo-ET 2D ('TubuleJ').

d'assemblages microtubulaires étudiés au sein de l'équipe TIPs (Umr 6026 Cnrs - Université de Rennes 1). Ils seront ensuite étendus à divers assemblages macro-moléculaires et aux cryo-sections de spécimens congelés-hydratés. Parallèlement, nous avons déjà proposé une méthode spectrale (analyse de Fourier) exploitant la redondance d'information pour la reconstruction 3D de microtubules et filaments (Blestel, Kervrann & Chrétien 2009, Blestel, Kervrann & Chrétien 2010b) (voir figure 5.1) à partir d'images 2D en projection.

Plus généralement, les biologistes ont désormais l'usage d'outils moléculaires (protéines fluorescentes normales ou mutées, exprimées dans des lignées cellulaires stables, approches des ARN, inférants), mécaniques (micro-patterning, microfluidiques), optiques (MSDC, Fast 4D-FRAP, photoactivation, photoablation, 4D + TIRF, Video-FLIM ...) ou structurels (Cryo-EM, Cryo-ET) très puissants pour normaliser ou perturber l'activité biologique dans des cellules, au niveau de l'ensemble du tissu et dans de petits organismes vivants (drosophile, ver, poisson-zèbre...). En combinant la quantité de données fournies en microscopie multi-dimensionnelle à haut débit avec des outils d'analyse automatique d'images, les biologistes collectent des données multi-paramétriques et hautement quantitatives de manière reproductible. La mise en correspondance de ces informations multi-modales et multi-échelles est désormais une exigence pour beaucoup de laboratoires de biologie cellulaire, mais un certain nombre d'obstacles techniques restent à surmonter. L'analyse de toutes ces informations devrait permettre ainsi de répondre à des questions fondamentales en biologie. Pour faire face, une démarche pluridisciplinaire et innovante couplant acquisition d'images en microscopie optique, modélisations statistiques, probabilistes et biophysique des processus multi-échelles, simulations *in silico* et expériences de biologie cellulaire

in vivo, est nécessaire.

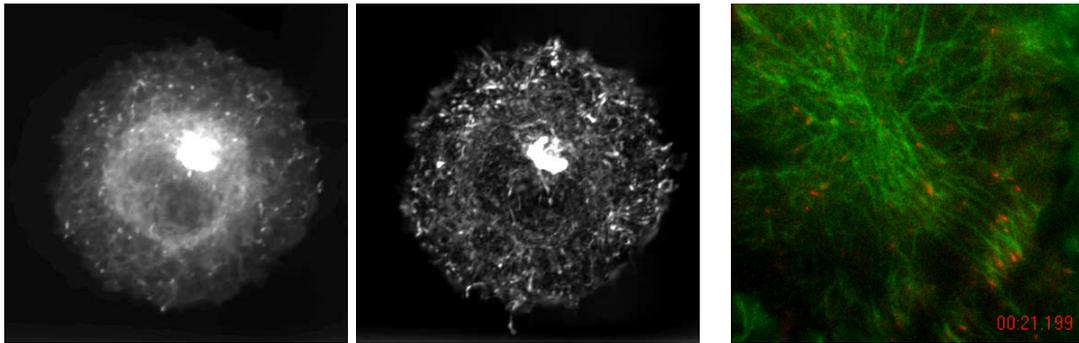
Dans mes perspectives, nous pouvons dégager plusieurs niveaux d'investigation dans le cadre d'études menées sur le système des endo-membranes, à l'échelle « cellule-unique » :

1. Un premier niveau visera à améliorer la résolution des images de microscopie optique à l'aide d'algorithmes numériques de super-résolution et de filtres non locaux, non-linéaires et plus élaborés que les filtres implémentés en imagerie super-résolution SIM ou PALM. Par ailleurs, nous serons également concernés par l'amélioration du rapport signal-sur-bruit, de plus en plus faible en pratique dès qu'il s'agit de préserver l'intégrité de la cellule (photo-toxicité vs. réduction du temps d'exposition). Des algorithmes de débruitage encore plus efficaces et plus rapides que ceux que nous avons déjà élaborés, issus de la théorie de l'apprentissage et spécialisés pour un contenu particulier, seront conçus conjointement avec nos collaborateurs experts en optique.
2. Un second niveau visera à extraire des informations des images et des vidéos pour quantifier les interactions dynamiques entre molécules ou groupes de molécules d'intérêt à l'échelle d'une cellule. Nous privilégierons les approches statistiques (méthodes non-paramétriques et bayésiennes) et variationnelles en veillant à intégrer les connaissances en biologie et en biophysique. Pour isoler, suivre et caractériser des centaines de petites entités (e.g. vésicules) dans des images bruitées, nous devons concevoir de nouveaux algorithmes d'estimation de trafic (modalité MSDC), d'estimation de mélanges de temps de vie de molécules en interaction (modalité FRET-FLIM), de détection avec un faible taux de fausses alarmes (modalité TIRF) ou de segmentation de polymères en microscopie électronique... ceci d'autant plus que les algorithmes usuels de vision par ordinateur ne sont pas vraiment adaptés à ces contextes nouveaux d'imagerie.
3. Le troisième niveau traitera de la cartographie de l'organisation spatiale et temporelle des édifices moléculaires étudiés. Les étapes intermédiaires auront pour but de mettre en correspondance les informations spatiales et temporelles obtenues par différentes modalités de microscopie optique (modalités MSDC, FLIM-FRET, TIRF, SIM...) et électronique (Vicidomini et al. 2008) et de modéliser les relations spatio-temporelles entre les événements dynamiques (voir Fig. 5.2). À cet égard, il semble assez naturel de prévoir des méthodes de détection de zone d'intérêt et d'adapter les grands principes de recalage d'images bien établis en imagerie médicale (Klein, Andersson, Ardekani, Ashburner, Avants, Chiang, Christensen, Collins, Gee, Hellier, Song, Jenkinson, Lepage, Rueckert, Thompson, Vercauteren, Woods, Mann & Parsey 2009, Roche, Pennec, Malandain & Ayache 2001, Vercauteren, Pennec, Perchant & Ayache 2009, Yeo, Sabuncu, Vercauteren, Holt, Amunts, Zilles, Golland & Fischl 2010, Yeo, Sabuncu, Vercauteren, Ayache, Fischl & Golland 2010). Les progrès en microscopie haute-résolution (PALM, SIM...) permettront de gérer les changements d'échelle, de l'ordre de quarante actuellement si on cherche à concilier déjà microscopie optique et microscopie électronique. L'étape finale, qui relève en fait de la « biologie des systèmes », consistera à établir les relations « structure-fonction » voire « fonction-structure » très prometteuses désormais, mettant en jeu des protéines/réseaux de protéines dans le système membranaire.
4. Le quatrième niveau relèvera davantage de la simulation et de la modélisation du trafic à l'aide d'algorithmes de « Tomographie de Réseau » ou de représentations graphiques.

L'intégration de connaissances biophysiques (e.g. dynamique stochastique des moteurs moléculaires) doit nous aider à mimer des comportements plus globaux, non seulement à l'échelle moléculaire mais aussi à des échelles supérieures (voies cellulaires, complexes macromoléculaires, cellule). Cette complexité devrait aussi conduire à des modélisations multi-échelles qui prennent en compte toutes ces informations dans un formalisme adéquat (e.g. équations différentielles stochastiques). Des expérimentations *in vivo* pour une classe large de *stimuli*, seront menées en parallèle : *i*) contraintes sur la forme de la cellule à l'aide de « micro-patrons » ; *ii*) inhibitions de moteurs moléculaires par SiRNA ; *iii*) perturbations par des drogues. Nous aborderons de ce fait le problème de la comparaison de dynamiques qui nécessite la définition de métriques géodésiques dans des espaces complexes.

Actuellement, les équipes de biologie cellulaire réclament des modèles précis capables de « déchiffrer » certains fonctionnements ou certains mécanismes cellulaires, comme ceux responsables du transport membranaire, voire de les prédire. Mettre en place une stratégie cohérente propre à atteindre ces objectifs nécessite dorénavant un rapprochement renforcé de nos compétences multidisciplinaires. Le défi ultime est de mettre en place un cadre de modélisation et une approche intégrative permettant d'évaluer l'influence de la mutation d'un gène sur la structure et la fonction d'une protéine, sur un processus ou un mécanisme cellulaire voire au-delà à l'échelle d'un tissu ou d'un organe. Cette ambition qui doit permettre de concevoir la cellule *in silico*, est très semblable à celle qui a motivé le projet Physiome (Hunter, Robbins & Noble 2002, Hunter & Borg 2003).

Microscopie optique (*in vivo*) : 200nm
(PICT IBiSA Umr 144 Institut Curie)

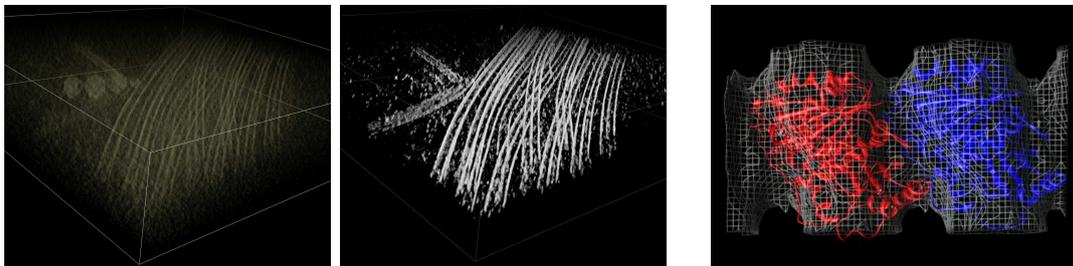


Rab6-GFP / réseau de microtubules
MSDC (300 ms/volume)

Tub-GFP / CLIP-170-mcherry
TIRF 2D (50 ms/image)

Microscopie super-résolution :
SIM (3D/100nm), PALM (2D/30nm)

Tomographie électronique (*in vitro*) : 5nm
(Umr 6026 Cnrs Université Rennes 1)



segmentation 3D de microtubules composés
de 13 protofilaments

alignement 3D de dimères
de tubuline sur un protofilament

FIG. 5.2. Analyse spatio-temporelle des microtubules : immobiles à l'échelle d'un trafic cellulaire (en haut à gauche), les microtubules sont des structures dynamiques qui polymérisent dès que leurs extrémités s'associent à la protéine CLIP-170 (en haut à droite). Ces interactions peuvent être analysées sur le plan structurel en microscopie électronique (en bas à droite). Chaque microtubule est composé de 13 ou 14 proto-filaments, eux-mêmes constitués de dimères de tubuline (en bas à droite).

- Aach, T., Dumbgen, L., Mester, R. & Toth, D. (2001), Bayesian illumination-invariant motion detection, *in* 'Proc. IEEE Int. Conf. on Image Process. (ICIP'01) (3)', Thessaloniki, Greece, pp. 640–643.
- Akaike, H. (1974), 'A new look at the statistical model identification', *IEEE Trans. on Automatic Control* **19**(6), 716–723.
- Akaike, H. (1978), 'A bayesian analysis of the minimum AIC procedure', *Ann. Inst. Statist. Math.* **30**, 9–14.
- Allende, H., Frery, A., Galbiati, J. & Pizarro, L. (2006), 'M-estimator with asymmetric influence function : the \mathcal{G}_a^0 distribution case', *J. Statist. Comput. and Simul.* **76**(11), 941–946.
- Alvarez, L., Gousseau, Y. & Morel, J. (1999), Scales in natural images and a consequence on their bounded variation, *in* 'Proc. Int. Conf. on Scale-Space Th. Comput. Vision (Scale-Space'99)', Kerkyra, Greece, pp. 247–258.
- Alvarez, L., Guichard, F., Lions, P. & Morel, J. (1993), 'Axioms and fundamental equations in image processing', *Arch. Rational Mech. Anal.* **16**(9), 200–257.
- Alvarez, L., Lions, P. & Morel, J. (1992), 'Image selective smoothing and edge detection by non-linear diffusion - 2', *SIAM J. Num. Anal.* **29**(3), 845–866.
- Arratia, R., Goldstein, L. & Gordon, L. (1989), 'Two moments suffice for poisson approximations : the Chen-Stein method', *Ann. Prob.* **17**(1), 9–25.
- Aujol, J.-F., Aubert, G., Blanc-Féraud, L. & Chambolle, A. (2005), 'Image decomposition into a bounded variation component and an oscillating component', *J. Math. Imaging and Vision* **22**, 71–88.
- Awate, S. & Whitacker, R. (2006), 'Unsupervised, information theoretic, adaptive image filtering for image restoration', *IEEE Trans. on Patt. Anal. and Mach. Intell.* **28**(3), 364–376.
- Axelrod, D. (1981), 'Cell-substrate contacts illuminated by total internal reflection fluorescence', *The J. of Cell Biol.* **89**, 141–145.

- Azencott, R. (1987), Image analysis and Markov fields, *in* ‘Proc. ICIAM SIAM Congress’, Paris, France.
- Azzabou, N. (2008), Variable bandwidth image models for texture-preserving enhancement of natural images, Thèse de doctorat, École des Ponts et Chaussées.
- Baddeley, A. & Van Lieshout, M. (1993), Stochastic geometry models in high-level vision, *in* K. Mardia & G. Kanji, eds, ‘Statistics and Images’, Vol. 20, Carfax Publishing, Abingdon, pp. 235–256.
- Ballester, C., Cubero-Castan, E., Gonzalez, M. & Morel, J. (2000), Contrast invariant image intersection, *in* ‘Adv. Math. Meth. in Measurement and Instruments’, p. 41–55.
- Barash, D. (2002), ‘A fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation’, *IEEE Trans. on Patt. Anal. and Mach. Intell.* **24**(6), 844–847.
- Barash, D. & Comaniciu, D. (2004), ‘A common framework for nonlinear diffusion, adaptive smoothing, bilateral filtering and mean shift’, *Image Vision Comput.* **22**(1), 73–81.
- Beaulieu, J. & Goldberg, M. (1989), ‘Hierarchy in picture segmentation : a stepwise optimization approach’, *IEEE Trans. on Patt. Anal. and Mach. Intell.* **11**(2), 150–163.
- Beaumont, M., Zhang, W. & Balding, D. (2002), ‘Approximate bayesian computation in population genetics’, *Genetics* **162**(4), 2025–2035.
- Bell, M. G. (1991), ‘The estimation of origine-destination matrices by constrained generalized least squares’, *Transportation Research B* **25**, 13–22.
- Benedek, C., Szirányi, T., Kato, Z. & Zerubia, J. (2007), A multi-layer MRF model for object-motion detection in unregistered airborne image-pairs, *in* ‘Proc. IEEE Int. Conf. on Image Process. (ICIP’07) (6)’, San Antonio, USA, pp. 141–144.
- Berger, A., Cabal, G., Fabre, E., Duong, T., Buc, H., Nehrbass, U., Olivo-Marin, J., Gadal, O. & Zimmer, C. (2008), ‘High-resolution statistical mapping reveals gene territories in live yeast’, *Nature Meth.* **5**, 1031 – 1037.
- Besag, J. (1974), ‘Spatial interaction and the statistical analysis of lattice systems’, *J. Royal Statist. Soc., Series B* (36), 192–236.
- Besag, J. (1975), ‘Statistical analysis of non-lattice data’, *The Statistician* **24**(3), 179–195.
- Besag, J. (1986), ‘On the statistical analysis of dirty pictures” (with discussions)’, *J. Royal Statist. Soc., Series B* (48), 259–302.
- Betzig, E., Patterson, G., Sougrat, R., Lindwasser, O., Olenych, S., Bonifacino, J., Davidson, M., Lippincott-Schwartz, J. & Hess, H. (2006), ‘Imaging intracellular fluorescent proteins at nanometer resolution’, *Science* **313**(5793), 1642–1645.
- Birgé, L. & Massart, P. (2001), ‘Gaussian model selection’, *J. Eur. Math. Soc.* **3**, 203–268.
- Black, M. & Anandan, P. (1996), ‘The robust estimation of multiple motions : parametric and piecewise-smooth flow fields’, *Comput. Vision and Image Understanding* **63**(1), 75–104.
- Black, M. & Fleet, D. (2000), ‘Probabilistic detection and tracking of motion boundaries’, *Int. J. Comput. Vision* **38**(3), 231–245.

- Black, M. & Rangarajan, A. (1996a), ‘On the unification of line processes, outlier rejection, and robust statistics with applications in early vision’, *Int. J. Comput. Vision* **19**(1), 57–92.
- Black, M. & Rangarajan, A. (1996b), ‘On the unification of line processes, outlier rejection, and robust statistics with applications in early vision’, *Int. J. Comput. Vision* **19**(1), 57–92.
- Black, M., Sapiro, G., Marimont, D. & Heeger, D. (1998a), ‘Robust anisotropic diffusion’, *IEEE Trans. on Image Process.* **7**(3), 421–432.
- Blake, A. & Zisserman, A. (1987), *Visual Reconstruction*, MIT Press, Cambridge, Mass.
- Bleil, J. & Bretscher, M. (1982), ‘Transferrin receptor and its recycling in hela cells’, *Eur. Mol. Biol. Org. J.* **1**(3), 351–355.
- Boiman, O. & Irani, M. (2005), Detecting irregularities in images and in video, in ‘Proc. IEEE Int. Conf. on Comput. Vision (ICCV’05) (1)’, Beijing, China, pp. 462–469.
- Bonneau, S., Dahan, M. & Cohen, L. D. (2005), ‘Single quantum dot tracking based on perceptual grouping using minimal paths in a spatiotemporal volume’, *IEEE Trans. on Image Process.* **14**(9), 1384–1395.
- Boyd, J. E., Meloche, J. & Vardi, Y. (1999), Statistical tracking in video traffic surveillance, in ‘Proc. IEEE Int. Conf. on Comput. Vision (ICCV’99) (1)’, Kerkyra, Greece, pp. 163–168.
- Boykov, Y. & Jolly, M. (2001), Interactive graph cuts for optimal boundary and region segmentation of objects in ND images, in ‘Proc. IEEE Int. Conf. on Comput. Vision (ICCV’01)’, Vancouver, Canada, pp. 105–112.
- Boykov, Y. & Kolmogorov, V. (2003), Computing geodesics and minimal surfaces via graph cuts, in ‘Proc. IEEE Int. Conf. on Comput. Vision (ICCV’03)’, Nice, France, pp. 26–33.
- Boykov, Y., Veksler, O. & Zabih, R. (2001), ‘Fast approximate energy minimization via graph cuts’, *IEEE Trans. on Patt. Anal. and Mach. Intell.* **23**(11), 1222–1239.
- Brox, T. & Cremers, D. (2009), ‘On local region models and a statistical interpretation of the piecewise smooth Mumford-Shah functional’, *Int. J. Comput. Vision* **84**(2), 184–193.
- Brox, T., Kleinschmidt, O. & Cremers, D. (2008), ‘Efficient nonlocal means for denoising of textural patterns’, *IEEE Trans. on Image Process.* **17**(7), 1083–1092.
- Brox, T. & Weickert, J. (2004), A tv flow based local scale measure for texture discrimination, in ‘Proc. Eur. Conf. on Comput. Vision (ECCV’04) (2)’, Prague, Czech Republic, pp. 578–590.
- Buades, A., Coll, B. & Morel, J. (2008), ‘Nonlocal image and movie denoising’, *Int. J. Comput. Vision* **76**(2), 123–139.
- Buades, A., Coll, B. & Morel, J. M. (2005), ‘A review of image denoising algorithms, with a new one’, *SIAM J. Multiscale Model. & Simul.* **4**(2), 490–530.
- Caselles, V., Coll, B. & Morel, J. (1999), ‘Topographic maps and local contrast changes in natural images’, *Int. J. Comput. Vision* **33**(1), 5–27.
- Caselles, V., Kimmel, R. & Sapiro, G. (1995), Geodesic active contours, in ‘Proc. IEEE Int. Conf. on Comput. Vision (ICCV’95)’, Boston, USA, pp. 694–699.
- Caselles, V., Kimmel, R. & Sapiro, G. (1997), ‘Geodesic active contours’, *Int. J. Comput. Vision* **22**(1), 61–79.

- Caselles, V., Morel, J., Sapiro, G. & Tannenbaum, A. (1998), ‘Special issue on partial differential equations and geometry-driven diffusion in image processing and analysis’, *IEEE Trans. on Image Process.* **7**(3), 269–473.
- Castro, R., Coates, M., Liang, G., Nowak, R. & Yu, B. (2004), ‘Network tomography : Recent developments’, *Statist. Science* **19**(3), 499–517.
- Catte, F., Lions, P., Morel, J. & T., C. (1992), ‘Image selective smoothing and edge detection by non-linear diffusion i’, *SIAM J. Num. Anal.* **29**(1), 182–193.
- Celeux, G., Forbes, F. & Peyrard, N. (2003), ‘Em procedures using mean field-like approximations for Markov model-based image segmentation’, *Patt. Recog.* **36**(1), 131–144.
- CGAL (2006), *CGAL-3.2 User and Reference Manual*, Editorial Board.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. & Prasher, D. (1994), ‘Green Fluorescent Protein as a marker for gene expression’, *Science* **263**, 802–805.
- Chalmond, B. (2000), *Éléments de modélisation pour l’analyse d’images*, Vol.33 of Mathématiques et Applications, Springer.
- Chambolle, A. (2004), ‘An algorithm for total variation minimization and applications’, *J. Math. Imaging and Vision* **20**(1-2), 89–97.
- Chan, T., Esedoglu, S. & Nikolova, M. (2006), ‘Algorithms for finding minimizers of image segmentation and denoising models’, *SIAM J. Appl. Math.* **66**(5), 1632–1648.
- Chan, T. F. & Vese, L. A. (2001), ‘Active contours without edges’, *IEEE Trans. on Image Process.* **10**(2), 266–277.
- Chavrier, P. & Goud, B. (1999), ‘The role of ARF and Rab GTPases in membrane transport’, *Cell Biol.* **11**, 466–475.
- Chen, S., Donoho, D. & Saunders, M. (2001), ‘Atomic decomposition by basis pursuit’, *SIAM Rev.* **43**(1), 129–159.
- Chenouard, N., Bloch, I. & Olivo-Marin, J.-C. (2009), Multiple hypothesis tracking in microscopy images, in ‘Proc. IEEE Int. Symp. on Biomedical Imaging : from nano to micro (ISBI’09)’, Boston, USA, pp. 1346–1349.
- Chesnaud, C., Réfrégier, P. & Boulet, V. (2000), ‘Statistical region snake-based segmentation adapted to different physical models’, *IEEE Trans. on Patt. Anal. and Mach. Intell.* **21**(11), 1145–1157.
- Cohen, I., Cohen, L. & Ayache, N. (1992), ‘Using deformable surfaces to segment 3-d images and infer differential structures’, *Comput. Vision, Graphics, and Image Process. : Image Understanding* **56**(2), 242–263.
- Cohen, L. (1991), ‘On active contour models and balloons’, *CVGIP : Image Understanding* **53**(2), 211–218.
- Cohen, L. & Kimmel, R. (1997), ‘Global minimum for active contour models : A minimal path approach’, *Int. J. Comput. Vision* **24**, 57–78.
- Cootes, T. F., Taylor, C. J., Cooper, D. H. & Graham, J. (1995), ‘Active shape models—their training and application’, *Comput. Vis. Image Underst.* **61**(1), 38–59.

- Cremers, D. and Tischhäuser, F., Weickert, J. & Schnörr, C. (2002), ‘Diffusion snakes : Introducing statistical shape knowledge into the Mumford–Shah functional’, *Int. J. Comput. Vision* **50**(3), 295–313.
- Cremers, D., Kohlberger, T. & Schnörr, C. (2003), ‘Shape statistics in kernel space for variational image segmentation’, *Patt. Recogn.* **36**(9), 1929–1943.
- Criminisi, A., Cross, G., Blake, A. & Kolmogorov, V. (2006), Bilayer segmentation of live video, in ‘Proc. IEEE Conf. on Comput. Vision Patt. Recogn. (CVPR’06) (1)’, Washington, USA, pp. 53–60.
- Criminisi, A., Pérez, P. & Toyama, K. (2004), ‘Region filling and object removal by exemplar-based image inpainting.’, *IEEE Trans. on Image Process.* **13**(9), 1200–1212.
- Dabov, K., Foi, A., Katkovnik, V. & Egiazarian, K. (2007), ‘Image denoising by sparse 3D transform-domain collaborative filtering’, *IEEE Trans. on Image Process.* **16**(8), 2080–2095.
- de Berg, M., van Kreveld, M., Overmars, M. & Schwarzkopf, O. (1997), *Computational geometry : algorithms and applications*, Springer-Verlag New York, Secaucus, NJ, USA.
- Deledalle, C., Denis, L. & Tupin, F. (2009), ‘Iterative weighted maximum likelihood denoising with probabilistic patch-based weights’, *IEEE Trans. on Image Process.* **18**(12), 2661–2672.
- Delpretti, S., Luisier, F., Ramani, S., Blu, T. & Unser, M. (2008), Multiframe SURE-let denoising of timelapse fluorescence microscopy images, in ‘Proc. IEEE Int. Symp. on Biomedical Imaging : from nano to macro (ISBI’08)’, Paris, France, pp. 149–152.
- Descombes, X. (2004), Méthodes stochastiques en analyse d’image : des champs de Markov aux processus ponctuels marqués, Habilitation à diriger des recherches, Université de Nice Sophia Antipolis.
- Desolneux, A., Moisan, L. & Morel, J.-M. (2000), ‘Meaningful alignments’, *Int. J. Comput. Vision* **40**(1), 7–23.
- Desolneux, A., Moisan, L. & Morel, J.-M. (2001), ‘Edge detection by Helmholtz principle’, *J. Math. Imaging Vis.* **14**(3), 271–284.
- Dibos, F. & Koepfler, G. (1998), Image denoising through a level set approach, in ‘Proc. IEEE Int. Conf. on Image Process. (ICIP’98) (3)’, Chicago, USA, pp. 264–268.
- Dijkstra, E. W. (1959), ‘A note on two problems in connexion with graphs’, *Numerische Mathematik* **1**(1), 269–271.
- Donoho, D. L. & Johnstone, I. M. (1994), ‘Ideal spatial adaptation by wavelet shrinkage’, *Biometrika* **81**(3), 425–455.
- Edelsbrunner, H. & Mücke, E. (1994), ‘Three-dimensional alpha shapes’, *ACM Trans. on Graphics* **13**, 43–72.
- Efros, A. & Leung, T. (1999), Texture synthesis by non-parametric sampling, in ‘Proc. IEEE Int. Conf. on Comput. Vision (ICCV’99) (2)’, Kerkyra, Greece, pp. 1033–1038.
- Elad, M. (2002), ‘On the origin of the bilateral filter and ways to improve it’, *IEEE Trans. on Image Process.* **11**(10), 1141–1151.
- Elad, M. & Aharon, M. (2006a), Image denoising via learned dictionaries and sparse representation, in ‘Proc. IEEE Conf. on Comput. Vision Patt. Recogn. (CVPR’06) (1)’, New York, USA, pp. 895–900.

- Elad, M. & Aharon, M. (2006c), ‘Image denoising via sparse and redundant representations over learned dictionary’, *IEEE Trans. on Image Process.* **15**(12), 3736–3745.
- Elgammal, A., Harwood, D. & Davis, L. (2000), Non-parametric model for background subtraction, in ‘Proc. Eur. Conf. on Comput. Vision (ECCV’00)’, Dublin, Ireland, pp. 751–767.
- Elmoataz, A., Lezoray, O. & Bougleux, S. (2008), ‘Nonlocal discrete regularization on weighted graphs : a framework for image and manifold processing’, *IEEE Trans. on Image Process.* **17**(7), 1047–1060.
- Embrechts, P., Kluppelberg, C. & Mikosch, T. (1997), *Modeling Extremal Events for Insurance and Finance*, Springer-Verlag.
- Fitzgibbon, A. W., Wexler, Y. & Zisserman, A. (2003), Image-based rendering using image-based priors, in ‘Proc. IEEE Int. Conf. on Comput. Vision (ICCV’03)’, Nice, France, pp. 1176–1183.
- Freeman, W. T., Pasztor, E. C. & Carmichael, O. T. (2000), ‘Learning low-level vision’, *Int. J. Comput. Vision* **40**(1), 25–47.
- Fukunaga, K. & Hostetler, L. (1975), ‘The estimation of the gradient of a density function, with applications in pattern recognition’, *IEEE Trans. on Inf. Theory* **21**(1), 32–40.
- Garcia-Parajo, M., Segers-Nolten, G., Veerman, J.-A., Greve, J. & van Hulst, N. (2000), ‘Real-time light-driven dynamics of the fluorescence emission in single Green Fluorescent Protein molecules’, *Proc. Nat. Acad. Sciences of USA* **97**(13), 7237–7242.
- Gasser, T., Sroka, L. & Jennen-Steinmetz, C. (1986), ‘Residual variance and residual pattern in nonlinear regression’, *Biometrika* **73**, 625–633.
- Geiger, D. & Yuille, A. (1991), ‘A common framework for image segmentation’, *Int. J. Comput. Vision* **6**(3), 1186–1191.
- Geman, D., Geman, S., Graffigne, C. & Dong, P. (1990), ‘Boundary detection by constrained optimization’, *IEEE Trans. on Patt. Anal. and Mach. Intell.* **12**(7), 609–628.
- Geman, S. & Geman, D. (1984), ‘Stochastic relaxation, gibbs distributions, and the bayesian restoration of images’, *IEEE Trans. on Patt. Anal. and Mach. Intell.* **6**(6), 721–741.
- Genovesio, A., Liedl, T., Emiliani, V., Parak, W. J., Coppey-Moisan, M. & Olivo-Marin, J.-C. (2006), ‘Multiple particle tracking in 3D+t microscopy : Method and application to the tracking of endocytosed quantum dots’, *IEEE Trans. on Image Process.* **15**(5), 1062–1070.
- Gilboa, G. & Osher, S. (2007), ‘Nonlocal linear image regularization and supervised segmentation’, *SIAM J. Multiscale Model. & Simul.* **6**(2), 595–630.
- Gilboa, G., Sochen, N. A. & Zeevi, Y. Y. (2006), ‘Variational denoising of partly textured images by spatially varying constraints’, *IEEE Trans. on Image Process.* **15**(8), 2281–2289.
- Girod, A., Storrie, B., Simpson, J., Johannes, L., Goud, B., Roberts, L., Lord, J., Nilson, T. & Pepperkok, R. (1999), ‘Evidence for a COP-I-independent transport route from the Golgi complex to the endoplasmic reticulum’, *Nature Cell Biol.* **1**, 423–430.
- Goldenshluger, A. & Nemirovski, A. (1997), ‘On spatial adaptive estimation of nonparametric regression’, *Math. Meth. Statist.* **6**(2), 135–170.

- Gordon, N., Salmond, D. & Smith, A. (1993), Novel approach to nonlinear/non-Gaussian Bayesian state estimation, in 'Proc. of IEEE', Vol. 140, pp. 107–113.
- Gousseau, Y. (2000), Distribution de formes dans les images naturelles, PhD thesis, Université de Paris 9.
- Grant, B. & Donaldson, J. (2005), 'Coupling between clathrin-coated-pit invagination, cortactin recruitment, and membrane scission observed in live cells', *Cell* **121**, 593–606.
- Grant, B. & Donaldson, J. (2009), 'Pathways and mechanisms of endocytic recycling', *Nat. Rev. Mol. Cell. Biol.* **10**, 597–608.
- Greig, D., Porteous, B. & Seheult, A. (1989), 'Exact maximum a posteriori estimation for binary images', *J. Royal Statist. Soc., Series B* (51), 271–279.
- Grenander, U., Chow, Y. & Keenan, D. M. (1991), *Hands : a Pattern Theoretic Study of Biological Shapes*, Springer-Verlag New York, Inc., New York, NY.
- Grenander, U. & Miller, M. (1994), 'Representations of knowledge in complex systems', *J. Royal Statist. Soc., series B* **56**(4), 549–603.
- Gustafsson, M., Shao, L., Carlton, P., Wang, C., Golubovskaya, I., Agard, D. & Sedat, J. (2008), 'Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination', *Biophysical J.* **94**(12), 4957–4970.
- Haussecker, H. & Fleet, D. (2001), 'Computing optical flow with physical models of brightness variation', *IEEE Trans. on Patt. Anal. and Mach. Intell.* **23**(6), 661–673.
- Heikkila, M. & Pietikainen, M. (2006), 'A textured-based method for modeling the background and detecting moving objects', *IEEE Trans. on Patt. Anal. and Mach. Intell.* **28**(4), 657–662.
- Heiler, M. & Schnörr, C. (2005), 'Natural image statistics for natural image segmentation', *Int. J. Comput. Vision* **63**(1), 5–19.
- Heitz, F. & Bouthemy, P. (1993), 'Multimodal estimation of discontinuous optical flow using Markov random fields', *IEEE Trans. on Patt. Anal. and Mach. Intell.* **15**(12), 1217–1232.
- Hirokawa, N. (1998), 'Kinesin and dynein superfamily proteins and the mechanism of organelle transport', *Science* **279**(5350), 519–526.
- Hoogenboom, J., van Dijk, E., Hernando, J., van Hulst, N. & Garcia-Parajo, M. (2005), 'Power-law-distributed dark states are the main pathway for photobleaching of single organic molecules', *Phys. Rev. Letters* **95**(9), 097401.1–097401.4.
- Huber, P. (1981), *Robust Statistics*, Wiley, New-York.
- Huet, S. (2002), Model selection for estimating the non zero component of a Gaussian vector, Technical Report 25, Université de Paris-Sud, Département de Mathématiques.
- Huet, S. (2006), 'Model selection for estimating the non zero components of a gaussian vector', *ESAIM Probab. Statist.* **10**, 164–183.
- Hunter, P. J. & Borg, T.K. (2003), 'Integration from protein to organs : the Physiome Project', *Nat. Rev. Mol. Cell Biol.*, **4**, 237–243.
- Hunter, P. J., Robbins, P. & Noble, D. (2002), 'The IUPS human Physiome Project', *Eur. J. Physiol.*, **445**(1), 1–9.

- Hurn, M. A., Husby, O. & Rue, H. (2003), A tutorial on image analysis, in J. Moller, ed., ‘Spatial Statistics and Computational Methods’, Vol. 173 of *Lecture Notes in Statist.*, Springer Verlag, Berlin, pp. 87–141.
- Ieng, S.-S., Tarel, J.-P. & Charbonnier, P. (2004), Evaluation of robust fitting based detection, in ‘Proc. Eur. Conf. on Comput. Vision (ECCV’04) (2)’, Prague, Czech Republic, pp. 341–352.
- Ince, S. & Konrad, J. (2005), Geometry-based estimation of occlusions from video frame pairs, in ‘Proc. IEEE Conf. on Acoust., Speech, and Signal Process. (ICASSP’05) (2)’, Philadelphia, USA, pp. 933–936.
- Ishikawa, H. (2003), ‘Exact optimization for Markov random fields’, *IEEE Trans. on Patt. Anal. and Mach. Intell.* **25**(10), 1333–1336.
- Jodoin, P.-M., Mignotte, M. & Konrad, J. (2007), ‘Statistical background subtraction using spatial cues’, *IEEE Trans. on Circ. and Syst. for Video Technol.* **17**(12), 1758–1763.
- Jodoin, P.-M., Mignotte, M. & Rosenberger, C. (2007), ‘Segmentation framework based on label field fusion’, *IEEE Trans. on Image Process.* **16**(10), 2535–2550.
- Jojic, N., Frey, B. J. & Kannan, A. (2003), Epitomic analysis of appearance and shape, in ‘Proc. IEEE Int. Conf. on Comput. Vision (ICCV’03)’, Nice, France, pp. 34–43.
- Juditsky, A. (1997), ‘Wavelet estimators : Adapting to unknown smoothness’, *Math. Meth. Statist.* (1), 1–20.
- Kasetkasem, T. & Varshney, P. (2002), ‘An image change detection algorithm based on Markov random field models’, *IEEE Trans. on Geoscience and Remote Sensing* **40**(8), 1815–1823.
- Kass, M., Witkin, A. & Terzopoulos, D. (1987), ‘Snakes : active contour models’, *Int J. Comput. Vision* **12**(1), 321–331.
- Katenka, N., Levina, E. & Michailidis, G. (2008), ‘Local vote decision fusion for target detection in wireless sensor networks’, *IEEE Trans. on Signal Process.* **56**(1), 329–338.
- Katkovnik, V., Egiazarian, K. & Astola, J. (2002), ‘Adaptive window size image denoising based on intersection of confidence intervals (ICI) rule’, *J. Math. Imaging and Vision* **16**(3), 223–235.
- Kindermann, S., Osher, S. & Jones, P. (2005), ‘Deblurring and denoising of images by nonlocal functionals’, *SIAM J. Multiscale Model. & Simul.* **4**(4), 1091–1115.
- Klein, A., Andersson, J., Ardekani, B., Ashburner, J., Avants, B., Chiang, M., Christensen, G., Collins, D. L., Gee, J., Hellier, P., Song, J., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R., Mann, J. J. & Parsey, R. (2009), ‘Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration’, *NeuroImage* **46**(3), 786–802.
- Koepfler, G., Lopez, C. & Morel, J. (1994), ‘A multiscale algorithm for image segmentation by variational method’, *SIAM J. Num. Anal.* **31**(1), 282–299.
- Kohli, P. & Torr, P. (2006), Measuring uncertainty in graph cut solutions : Efficiently computing min-marginal energies using dynamic graph cuts, in ‘Proc. Eur. Conf. on Comput. Vision (ECCV’06) (2)’, Graz, Austria, pp. 30–43.

- Kolmogorov, V. & Zabih, R. (2001), Computing visual correspondence with occlusions via graph cuts, in 'Proc. IEEE Int. Conf. on Comput. Vision (ICCV'01)', Vancouver, Canada, pp. 508–515.
- Koloydenko, A. & Geman, D. (2006), Ordinal coding of image microstructure, in 'IPCV', pp. 613–620.
- Kumar, M., Torr, P. & Zisserman, A. (2005), Obj cut, in "Proc. IEEE Conf. on Comput. Vision Patt. Recog. (CVPR'05)", Vol. 1, San Diego, USA, pp. 18–25.
- Kumar, S. & Hebert, M. (2006), 'Discriminative random fields', *Int. J. Comput. Vision* **68**(2), 179–201.
- Lafferty, J., McCallum, A. & Pereira, F. (2001), Conditional random fields : Probabilistic models for segmenting and labeling sequence data, in 'Proc. Int. Conf. on Mach. Learning (ICML'01)', p. 282–289.
- Le Pennec, E. & Mallat, S. (2005), 'Sparse geometric image representations with bandelets', *IEEE Trans. on Image Process.* **14**(4), 423–438.
- Lebarbier, E. (2005), 'Detecting multiple change-points in the mean of gaussian process by model selection', *Signal Process.* **85**(4), 717–736.
- Leclerc, Y. (1989), 'Constructing simple stable descriptions for image partitioning', *Int J. Comput. Vision* **3**, 73–102.
- Lee, A. B., Pedersen, K. S. & Mumford, D. (2003), 'The nonlinear statistics of high-contrast patches in natural images', *Int. J. Comput. Vision* **54**(1-3), 83–103.
- Lee, J. (1983), 'Digital image smoothing and the sigma filter', *Comput. Vision, Graphics, and Image Process.* **24**, 255–269.
- Lellmann, J., Becker, F. & Schnörr, C. (2009), Convex optimization for multi-class image labeling with a novel family of total variation based regularizers, in 'Proc. IEEE Int. Conf. on Comput. Vision (ICCV'09)', Kyoto, Japan, pp. 646–653.
- Lepski, O. (1991), 'Asymptotically minimax adaptive estimation 1 : upper bounds', *Theory Prob. and Appl.* **36**(4), 654–659.
- Lepski, O., Mammen, E. & Spokoiny, V. (1997), 'Optimal spatial adaptation to inhomogeneous smoothness : an approach based on kernel estimates with variable bandwidth selectors', *Ann. Statist.* **25**(3), 929–947.
- Leung, T. & Malik, J. (1996), Detecting, localizing and grouping repeated scene elements from an image, in 'Proc. Eur. Conf. on Comput. Vision (ECCV'96) (1)', Cambridge, UK, pp. 546–555.
- Levina, E. & Bickel, P. (2006), 'Texture synthesis and non-parametric resampling of random fields', *Ann. Statist.* **34**(4), 1751–1773.
- Lindeberg, T. (1998), 'Edge detection and ridge detection with automatic scale selection', *Int. J. of Comput. Vision* **30**(2), 117–156.
- Lisani, J. & Morel, J.-M. (2003), Detection of major changes in satellite images, in 'Proc. IEEE Int. Conf. on Image Process. (ICIP'03) (1)', Barcelona, Spain, pp. 941–944.

- Louchet, C. (2008), Modèles variationnels et bayésiens pour le débruitage d'images : de la variation totale vers les moyennes non-locales, Thèse de doctorat, Université de Paris Descartes.
- Louchet, C. & Moisan, L. (2008), Total variation denoising using posterior expectation, in 'Proc. Eur. Signal Processing Conf. (EUSIPCO'08) (3)', Lausanne, Switzerland, pp. 1741–1744.
- Mairal, J., Bach, F., Ponce, J., Sapiro, G. & A., Z. (2009), Non-local sparse models for image restoration, in 'Proc. IEEE Int. Conf. on Comput. Vision (ICCV'09)', Kyoto, Japan, pp. 2272–2279.
- Mairal, J., Sapiro, G. & Elad, M. (2008), 'Learning multiscale sparse representation for image and video restoration', *SIAM J. Multiscale Model. & Simul.* **7**(1), 214–241.
- Malladi, M. & Sethian, J. (1996), 'Image processing : Flows under min/max curvature and mean curvature', *Graphical Models and Image Process.* **58**(2), 127–141.
- Mallat, S. (1999), *A Wavelet Tour of Signal Processing*, Academic press.
- Mallows, C. (1974), 'Some comments on c_p ', *Technometrics* **15**, 661–675.
- Marroquin, J.L. and Mitter, S. & Poggio, T. (1987), 'Probabilistic solution of ill-posed problems in computational vision', *J. Amer. Statist. Assoc.* **82**, 76–89.
- Matheron, G. (1975), *Random Sets and Integral Geometry*, John Wiley, New York.
- Møller, J. & Waagepetersen, R. (1998), 'Markov connected component fields', *Adv. in Appl. Prob.* **30**, 1–35.
- Monasse, P. & Guichard, F. (2000), 'Fast computation of a contrast-invariant image representation', *IEEE Trans. on Image Process.* **9**(5), 860–872.
- Morel, J. & Solimini, S. (1994), *Variational Methods in Image Segmentation*, Birkhauser.
- Morrison, J. & Smith, J. (2002), 'Stochastic modeling of flood peaks using the generalized extreme value distribution', *Water Resources Research* **38**(12), 41.1–41.12.
- Mrázek, P., Weickert, J. & Bruh, A. (2006), On robust estimation and smoothing with spatial and tonal kernel', *Geometric Properties from Incomplete Data*, pp. 335–352, Springer.
- Mumford, D. (1994), The bayesian rationale for energy functionals, in B. Romeny, ed., 'Geometry-Driven Diffusion in Comput. Vision', Kluwer Academic, pp. 141–153.
- Mumford, D. & Shah, J. (1989), 'Optimal approximations by piecewise smooth functions and variational problems', *Communication on Pure and Appl. Math.* **42**(5), 577–685.
- Murtagh, F., Starck, J. & Bijaoui, A. (1995), 'Image restoration with noise suppression using a multiresolution support', *Astronomy and Astrophysics* **112**, 197–189.
- Negahdaripour, S. (1998), 'Revised definition of optical flow : integration of radiometric and geometric clues for dynamic scene analysis', *IEEE Trans. on Patt. Anal. and Mach. Intell.* **20**(9), 961–979.
- Nemsick, L. & Geraniotis, E. (1992), 'Adaptive multichannel detection of frequency-hopping signals', *IEEE Trans. on Communications* **40**(9), 1502–1511.
- Neumann, B., Held, M., Liebel, U., Erfle, H., Rogers, P., Pepperkok, R. & Ellenberg, J. (2006), 'High-throughput rna screening by time-lapse imaging of live human cells', *Nature Meth.* **3**, 385–390.

- Niu, R., Varshney, P., Moore, M. & Klamer, D. (2004), Decision fusion in a wireless sensor network with a large number of sensors, in ‘Proc. Int. Conf. Inf. Fusion (FUSION’04)’, Stockholm, Sweden, pp. 1–8.
- Nordström, K. (1990), ‘Biased anisotropic diffusion : a unified regularization and diffusion approach to edge detection’, *Image and Vision Comput.* **8**(4), 318–327.
- Odobez, J.-M. & Bouthemy, P. (1995), ‘Robust multiresolution estimation of parametric motion models’, *J. Vis. Comm. Image Repres.* **6**(4), 348–365.
- Osher, S. & Sethian, J. (1988), ‘Fronts propagating with curvature dependent speed : algorithms based on the hamilton-jacobi formulation’, *J. Comput. Physics* **79**, 12–49.
- Osher, S., Sole, A. & Vese, L. (2002), ‘Image decomposition and restoration using total variation minimization and the h^{-1} norm’, *SIAM J. Multiscale Model. & Simul.* **1**(3), 349–370.
- O’Sullivan, F. & Qian, M. (1994), ‘A regularized contrast statistic for object boundary estimation – implementation and statistical evaluation’, *IEEE Trans. on Patt. Anal. and Mach. Intell.* **16**(6), 561–570.
- Paragios, N. & Deriche, R. (2000), Coupled geodesic active regions for image segmentation : a level set approach, in ‘Eur. Conf. on Comput. Vis. (ECCV’00) (2)’, Dublin, Ireland, pp. 224–240.
- Pepperkok, R. & Ellenberg, J. (2006), ‘High-throughput fluorescence microscopy for systems biology’, *Nature Rev. Mol. Cell Biol.* **7**, 690–696.
- Perona, P. & Malik, J. (1990a), ‘Scale-space and edge detection using anisotropic diffusion’, *IEEE Trans. on Patt. Anal. and Mach. Intell.* **12**(7), 629–639.
- Perona, P. & Malik, J. (1990b), ‘Scale-space and edge detection using anisotropic diffusion’, *IEEE Trans. on Patt. Anal. and Mach. Intell.* **12**(7), 629–639.
- Pilet, J., Strecha, C. & Fua, P. (2008), Making background subtraction robust to sudden illumination changes, in ‘Proc. Eur. Conf. on Comput. Vision (ECCV’08) (4)’, Marseille, France, pp. 567–580.
- Pizurica, A. & Philips, W. (2006), ‘Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising’, *IEEE Trans. on Image Process.* **15**(3), 654–665.
- Pock, T., Chambolle, A., Cremers, D. & Bischof, H. (2009), A convex relaxation approach for computing minimal partitions, in ‘Proc. IEEE Conf. on Comput. Vision Patt. Recog. (CV-PR’09)’, Miami, USA, pp. 810–817.
- Polzehl, J. & Spokoiny, V. (1999), Vector adaptive weights smoothing with application to MRI, Technical Report 519, Weirstrass Institute for Applied Analysis and Statistics, Berlin.
- Polzehl, J. & Spokoiny, V. (2000), ‘Adaptive weights smoothing with applications to image restoration’, *J. the Royal Statist. Soc. : Series B* **62**(2), 335–354.
- Portilla, J., Strela, V., Wainw. & Simoncelli, E. (2003), ‘Image denoising using scale mixtures of Gaussians in the wavelet domain’, *IEEE Trans. on Image Process.* **12**(11), 1338–1351.
- Pérez, P. (1998), ‘Markov random fields and images’, *CWI Quarterly* **11**(4), 413–437.

- Racine, V. (2006), Quantification des Dynamiques Cellulaires par Analyse de Données de Vidéo-Microscopie 3D+T, PhD thesis, Université Pierre et Marie Curie - Paris 6.
- Racine, V., Saschse, M., Salamero, J., Fraissier, V., Trubuil, A. & Sibarita, J. (2007), ‘Visualization and quantification of vesicle trafficking on a 3D cytoskeleton network in living cells’, *J. Microscopy* pp. 214–228.
- Radke, R., Andra, S., Al Kofahi, O. & Roysam, B. (2005), ‘Image change detection algorithms : a systematic survey’, *IEEE Trans. on Image Process.* **14**(3), 294–307.
- Ramani, S., Vonesch, C. & Unser, M. (2008), Deconvolution of 3d micrographes with automatic risk minimization, in ‘Proc. IEEE Int. Symp. on Biomedical Imaging : from nano to macro (ISBI’08)’, Paris, France, pp. 732–735.
- Rissanen, J. (1978), ‘Modeling by shortest data description’, *Automatica* **14**, 465–471.
- Roche, A., Pennec, X., Malandain, G. & Ayache, N. (2001), ‘Rigid registration of 3D ultrasound with MR images : a new approach combining intensity and gradient information’, *IEEE Trans. on Medical Imaging* **20**(10), 1038–1049. PMID : 11686439.
- Rosin, P. (2002), ‘Thresholding for change detection’, *Comput. Vis. Image Underst.* **86**(2), 79–95.
- Roth, S. & Black, M. (2005), Fields of experts : A framework for learning image priors, in ‘Proc. IEEE Conf. on Comput. Vision Patt. Recog. (CVPR’05)’, San Diego, USA, pp. 860–867.
- Ruckstuhl, A., Jacobson, M., Field, R. & Dodd, J. (2001), ‘Baseline subtraction using robust local regression estimation’, *J. Quantitative Spectroscopy and Radiative Transfer* **68**(2), 179–193.
- Rudin, L., Osher, S. & Fatemi, E. (1992), ‘Nonlinear Total Variation based noise removal algorithms’, *Physica D* **60**(1-4), 259–268.
- Sabater, N., Almansa, A. & Morel, J.-M. (2008), Rejecting wrong matches in stereovision, Technical Report 2008-28, CMLA.
- Sage, D., Neumann, F., Hediger, H., Gasser, S. & Unser, M. (2005), ‘Automatic tracking of individual fluorescence particles : Application to the study of chromosome dynamics’, *IEEE Trans. on Image Process.* **14**(9), 1372–1383.
- Salamero, J. (2008), ‘La "révolution verte" est en marche’, *Médecine/Sciences* **24**, 987–988.
- Salembier, P. & Serra, J. (1995), ‘Flat zones filtering, connected operators, and filters by reconstruction’, *IEEE Trans. on Image Process.* **4**(8), 1153–1160.
- Saporta, G. (1978), *Théories et Méthodes de la Statistique*, Technip.
- Sbalzarini, I. & Koumoutsakos, P. (2005), ‘Feature point tracking and trajectory analysis for video imaging in cell biology’, *J. Struct. Biol.* **151**, 182–195.
- Schnörr, C. (1998), ‘A study of a convex variational diffusion approach for image segmentation and feature extraction’, *J. of Math. Imag. and Vision* **8**(3), 271–292.
- Schwarz, G. (1978a), ‘Estimating the dimension of a model’, *Ann. Statist.* **6**, 461–464.
- Schölkopf, B., Tsuda, K. & Vert, J.-P. (2004), *Kernel Methods in Computational Biology*, MIT Press.
- Seabra, M., Mules, E. & Hume, A. (2002), ‘Rab GTPase, intracellular traffic and disease’, *Trends Mol. Med.* **8**(1), 23–30.

- Serra, J. (1982), *Image Analysis and Mathematical Morphology*, Academic Press.
- Sethian, J. (1996), *Level Sets Methods : Evolving Interfaces in Geometry, Fluid Mechanics, Comput. Vision, and Material Science*, Cambridge University Press.
- Shechtman, E. & Irani, M. (2007), Matching local self-similarities across images and videos, in “Proc. IEEE Conf. on Comput. Vision Patt. Recog. (CVPR’07)”, Minneapolis, USA, pp. 1–8.
- Shi, J. & Thomasi, C. (1994), Good feature to track, in ‘Proc. IEEE Conf. on Comput. Vision and Patt. Recog. (CVPR’94)’, Seattle, USA, pp. 593–600.
- Shimomura, O., Johnson, F. & Saiga, Y. (1962), ‘Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, aequorea”, *J. Cellular and Comparative Physiology* **59**(3), 223–239.
- Sibarita, J.-B., Magnin, H. & De Mey, J. R. (2002), Ultra-fast 4D microscopy and high throughput distributed deconvolution, in ‘Proc. IEEE Int. Symp. on Biomedical Imaging : from nano to macro (ISBI’02)’, Washington, USA, pp. 769–772.
- Smal, I., Draegestein, K., Galjart, N., Niessen, W. & Meijering, E. (2008), ‘Particle filtering for multiple object tracking in dynamic fluorescence microscopy images : Application to microtubule growth analysis’, *IEEE Trans. on Medical Imaging* **27**(6), 789–804.
- Smal, I., Meijering, E., Draegestein, K., Galjart, N., Grigoriev, I., Akhmanova, A., van Royen, M., Houtsmuller, A. & Niessen, W. (2008), ‘Multiple object tracking in molecular bioimaging by Rao-Blackwellized marginal particle filtering’, *Medical Image Anal.* **12**(6), 764–777.
- Smith, S. M. & Brady, J. M. (1997), ‘Susan - a new approach to low level image processing’, *Int. J. Comput. Vision* **23**(1), 45–78.
- Sochen, N. A., Kimmel, R. & Bruckstein, A. M. (2001), ‘Diffusions and confusions in signal and image processing’, *J. Math. Imaging and Vision* **14**(3), 195–209.
- Stankovic, L. (2004), ‘Performance analysis of the adaptive algorithm for bias-to-variance tradeoff’, *IEEE Trans. on Signal Process.* **52**(5), 1228–1234.
- Starck, J.-L., Candes, E. & Donoho, D. L. (2000), ‘The curvelet transform for image denoising’, *IEEE Trans. on Image Process.* **11**(6), 670–684.
- Starck, J.-L. & Murtagh, F. (1998), ‘Automatic noise estimation from the multiresolution support’, *Publications of the Astr. Soc. of the Pacific* **110**, 193–199.
- Stauffer, C. & Grimson, W. (1999), Adaptive background mixture models for real time tracking, in “Proc. IEEE Conf. on Comput. Vision Patt. Recog. (CVPR’99)”, Ft. Collins, USA, pp. 246–252.
- Sun, J., Li, Y., Kang, S.-B. & Shum, H.-Y. (2005), Symmetric stereo matching for occlusion handling, in “Proc. IEEE Conf. on Comput. Vision Patt. Recog. (CVPR’05)”, San Diego, USA, pp. 299–406.
- Sun, J., Zhang, W., Tang, X. & Shum, H.-Y. (2006), Background cut, in ‘Proc. Eur. Conf. on Comput. Vision (ECCV’06) (2)’, Graz, Austria, pp. 628–641.
- Szummer, M., Kohli, P. & Hoiem, D. (2008), Learning crfs using graph cuts, in ‘Proc. Eur. Conf. on Comput. Vision (ECCV’08) (2)’, Marseille, France, pp. 582–595.

- Tamin, O. & Willumsen, L. (1989), ‘Transport demand model estimation from traffic counts’, *Transportation Research* **16**(1), 3–26.
- Thery, M., Racine, V., Pepin, A., Piel, M., Chen, Y., Sibarita, J.-B. & Bornens, M. (2005), ‘The extracellular matrix guides the orientation of the cell division axis’, *Nature Cell Biol.* **7**, 947–953.
- Thomann, D., Dorn, J., Sorger, P. & Danuser, G. (2003), ‘Automatic fluorescent tag localization ii : improvement in super-resolution by relative tracking’, *J. Microscopy* **211**(3), 230–248.
- Thomann, D., Rines, D. ., Sorger, P. & Danuser, G. (2002), ‘Automatic fluorescent tag detection in 3D with super-resolution : application to the analysis of chromosome movement’, *J. Microscopy* **208**.
- Tikhonov, A. & Arsenin, V. (1977), *Solution of Ill-posed Problems*, Winston & Sons, Washington.
- Tomasi, C. & Manduchi, R. (1998), Bilateral filtering for gray and color images, in ‘Proc. IEEE Int. Conf. on Comput. Vision (ICCV’98)’, Bombay, India, pp. 839–846.
- Toyama, K., Krumm, J., Brumitt, B. & Meyers, B. (1999), Wallflower : principles and practice of background maintenance, in ‘Proc. IEEE Conf. on Comput. Vision Patt. Recog. (CVPR’99) (1)’, Ft. Collins, USA, pp. 255–261.
- Tschumperlé, D. (2006), ‘Fast anisotropic smoothing of multi-valued images using curvature-preserving pde’s’, *Int. J. Comput. Vision* **68**(1), 65–82.
- Tschumperlé, D. & Deriche, R. (2005), ‘Vector-valued image regularization with PDE’s : a common framework for different applications’, *IEEE Trans. on Patt. Anal. and Mach. Intell.* **27**(4), 506–517.
- Tu, Z. & Zhu, S.-C. (2002), ‘Image segmentation by data-driven Markov chain Monte Carlo’, *IEEE Trans. on Patt. Anal. and Mach. Intell.* **24**(5), 657–673.
- Uzan-Gafsou, S., Bausinger, H., Proamer, F., Monier, S., D., L., Cazenave, J., Goud, B., de la Salle, H., Hanau, D. & Salamero, J. (2007), ‘Rab11a controls the biogenesis of birbeck granules by regulating langerin recycling and stability’, *Mol. Biol. of the Cell* **18**, 3169–3179.
- Valladeau, J., Ravel, O., Dezutter-Dambuyant, C., Moore, K., Kleijmeer, M., Liu, Y., Duvert-Frances, V., Vincent, C., Schmitt, D., Davoust, J., Caux, C., Lebecque, S. & Saeland, S. (2000), ‘Langerin, a novel c-type lectin specific to langerhans cells, is an endocytic receptor that induces the formation of birbeck granules’, *Immunity* **12**, 71–81.
- van de Weijer, J. & van den Boomgaard, R. (2001), Local mode filtering, in ‘Proc. IEEE Conf. on Comput. Vision Patt. Recog. (CVPR’01) (2)’, Kauai, USA, pp. 428–433.
- Vardi, Y. (1996), ‘Network tomography : Estimation source-destination traffic intensities from link data’, *J. Amer. Statist. Ass.* **91**(433), 365–377.
- Vercauteren, T., Pennec, X., Perchant, A., & Ayache, N. (2009), ‘Diffeomorphic Demons : Efficient Non-parametric Image Registration’, *NeuroImage* **45**(1, Supp.1), S61–S72.
- Vese, L. (1999), ‘A method to convexify functions via curve evolution’, *Commun. Partial Differential Equations* **24**(9-10), 1573–1591.

- Vicidomini, G., Gagliani, M., Canfora, M., Cortese, K., Frosi, F., Santangelo, C., Di Fiore, P., Boccacci, P., Diaspro, A. & Tacchetti, C. (2008), 'Automated 3d correlative light-electron microscopy method', *Traffic* **9**, 1828–1838.
- Vincent, L. & Soille, P. (1991), 'Watershed in digital spaces : an efficient algorithm based on immersion simulations', *IEEE Trans. on Patt. Anal. and Mach. Intell.* **13**(6), 583–598.
- Viola, P. & Jones, M. (2001), Rapid object detection using a boosted cascade of simple features, in 'Proc. IEEE Conf. on Comput. Vision and Patt. Recog. (CVPR'01) (1)', Hawai, USA, pp. 511–518.
- Wang, Y. & Zhu, S. C. (2002), A generative method for textured motion : Analysis and synthesis, in 'Proc. Eur. Conf. on Comput. Vision (ECCV'02)', Copenhagen, Denmark, pp. 583–598.
- Weickert, J. (1999), 'Coherence-enhancing diffusion filtering', *Int. J. Comput. Vision* **31**(2-3), 111–127.
- White, J., Johannes, L., Mallar, F., Girod, A., Stephan, G., Reinsh, S., Keller, P., Tzschaschel, B., Echard, A., Goud, B. & Stelzer, E. (1999), 'Rab6 coordinates a novel golgi to ER retrograde transport pathway in live cells', *The J. Cell Biol.* **147**(4), 743–760.
- Winkler, G. (1995), *Image analysis, random fields and dynamic Monte Carlo methods*, Springer.
- Yeo, B.T.T., Sabuncu, M., Vercauteren, T., Holt, D., Amunts, K., Zilles, K., Golland, P. & Fischl, B. (2010), 'Learning Task-Optimal Registration Cost Functions for Localizing Cytoarchitecture and Function in the Cerebral Cortex', *IEEE Transactions on Medical Imaging*. (In Press)
- Yeo, B.T.T., Sabuncu, M., Vercauteren, T., Ayache, N., Fischl, B. & Golland, P. (2010), 'Spherical Demons : Fast Diffeomorphic Landmark-Free Surface Registration', *IEEE Transactions on Medical Imaging* **29**(3), 650-668.
- Yezzi, A., Tsai, A. & Willsky, A. (1999), A statistical approach to snakes for bimodal and trimodal imagery, in 'Int. Conf. on Comput. Vis.', Kerkyra, Greece, pp. 898–903.
- Younes, L. (2000), Calibrating parameters of cost functionals, in 'Eur. Conf. on Comput. Vis. (ECCV'00) (2)', Dublin, Ireland, pp. 212–223.
- Zhang, B., Fadili, M. & Starck, J.-L. (2006), Multi-scale variance stabilizing transform for multi-dimensional Poisson count image denoising, in 'Proc. IEEE Int. Conf. on Acoust., Speech, and Signal Process. (ICASSP'06) (2)', Toulouse, France, pp. 81–84.
- Zhu, S.-C. (1999), 'Stochastic jump-diffusion process for computing medial axes in Markov random fields', *IEEE Trans. on Patt. Anal. and Mach. Intell.* **21**(11), 1158–1169.
- Zhu, S. & Mumford, D. (1997), 'Prior learning and gibbs reaction-diffusion', *IEEE Trans. on Patt. Anal. and Mach. Intell.* **19**(11), 1236–1250.
- Zhu, S. & Yuille, A. (1996), 'Region competition : unifying snakes, region growing, and bayes/MDL for multiband image segmentation', *IEEE Trans. on Patt. Anal. and Mach. Intell.* **18**(9), 884–900.

Bibliographie

Publications de l'auteur référencées dans le mémoire

- Blestel, S., Kervrann, C. & Chrétien, D. (2010a), 'Electron cryo-tomography of tubulin assemblies : automatic segmentation using shape recognition', to be submitted to *J. Struct. Biol.*
- Blestel, S., Kervrann, C. & Chrétien, D. (2010b), 'TubulinJ : a toolbox to process electron cryo-microscope images of vitrified microtubules', to be submitted to *J. Struct. Biol.*
- Blestel, S., Kervrann, C. & Chrétien, D. (2009), A Fourier-based method for detecting curved microtubule centers : application to straightening of cryo-electron microscope images, *in* 'Proc. IEEE Int. Symp. on Biomedical Imaging : from nano to macro (ISBI'09)', Boston, USA, pp. 298-301.
- Boulanger, J., Gidon, A., Kervrann, C. & J. Salamero, J. (2009), 'A patch-based method for repetitive and transient event detection in fluorescence imaging', *PLoS ONE*. (in revision)
- Boulanger, J., Kervrann, C. & Bouthemy, P. (2005a), Adaptive spatio-temporal restoration for 4d fluorescence microscopic imaging, *in* 'Proc. Int. Conf. on Medical Image Comput. and Comput. Assisted Intervention (MICCAI'05)', Palm Springs, USA, pp. 893-901.
- Boulanger, J., Kervrann, C. & Bouthemy, P. (2005b), An adaptive statistical method for denoising 4d fluorescence image sequences with preservation of spatio-temporal discontinuities, *in* 'Proc. IEEE Int. Conf. on Image Process. (ICIP'05)', Genova, Italy, pp. 145-148.
- Boulanger, J., Kervrann, C. & Bouthemy, P. (2006), Estimation of dynamic background for fluorescence video-microscopy, *in* 'Proc. IEEE Int. Conf. on Image Process. (ICIP'06)', Atlanta, USA, pp. 2509-2512.
- Boulanger, J., Kervrann, C. & Bouthemy, P. (2007), 'Space-time adaptation for patch based image sequence restoration', *IEEE Trans. on Patt. Anal. and Mach. Intell.* **29**(6), 1096-1102.

- Boulanger, J., Kervrann, C. & Bouthemy, P. (2009), ‘A simulation and estimation framework for intracellular dynamics and trafficking in video-microscopy and fluorescence imagery’, *Medical Image Anal.* **13**, 132–142.
- Boulanger, J., Kervrann, C., Bouthemy, P., Elbau, P., Sibarita, J.-B. & Salamero, J. (2010), ‘Patch-based non-local functional for denoising fluorescence microscopy image sequences’, *IEEE Trans. on Medical Imaging* **29**(2), 442–454.
- Boulanger, J., Sibarita, J.-B., Kervrann, C. & Bouthemy, P. (2008), Non-parametric regression for patch-based fluorescence microscopy image sequence denoising, *in* ‘Proc. IEEE Int. Symp. on Biomedical Imaging : from nano to macro (ISBI’08)’, Paris, France, pp. 748–751.
- Carlton, P., Boulanger, J., Kervrann, C., Sibarita, J.-B., Salamero, J., Gordon-Messer, S., Haber, J., Haase, S., Shao, L., Winoto, L., Matsuda, A., Kner, P., Usawa, S., Strukov, Y., Gustafsson, M., Kam, Z., Agard, D. & Sedat, J. (2010), ‘Fast live simultaneous multi-wavelength 4-dimensional optical microscopy’, Proceedings of the National Academy of Sciences USA. (to appear) Science.
- Chessel, A., Cinquin, B., Bardin, S., Salamero, J. & Kervrann, C. (2009), Computational geometry-based scale-space and modal image decomposition : application to light video-microscopy imaging, *in* ‘Proc. Conf. on Scale Space and Variational Meth. (SSVM’09)’, Voss, Norway, pp. 770–781.
- Chessel, A., Cinquin, B., Boulanger, J., Salamero, J. & Kervrann, C. (2010), A detection-based framework for the analysis of recycling in tirf microscopy, *in* ‘Proc. IEEE Int. Symp. on Biomedical Imaging : from nano to macro (ISBI’10)’, Rotterdam, The Netherlands.
- Chessel, A., Fablet, R., Cao, F. & Kervrann, C. (2006), Orientation interpolation and applications, *in* ‘Proc. IEEE Int. Conf. on Image Process. (ICIP’06)’, Atlanta, USA, pp. 1561–1564.
- Chessel, A., Fablet, R., Kervrann, C. & Cao, F. (2008), Otolith image analysis by computer vision : extracting growth rings and recovering shape evolution of accretionary structures, *in* ‘Proc. Int. Conf. on Bio-inspired Systems and Signal Process. (BIOSIGNALS’08) (2)’, Funchal, Portugal, pp. 490–497.
- Chessel, A., Pécot, T., Bardin, S., Kervrann, C. & Salamero, J. (2009), Evaluation of image sequences additive decomposition algorithms for membrane analysis in fluorescence video-microscopy, *in* ‘Proc. IEEE Int. Symp. on Biomedical Imaging : from nano to macro (ISBI’09)’, Boston, USA, pp. 1099–1102.
- Coupé, P., Hellier, P., Kervrann, C. & Barillot, C. (2009), ‘Nonlocal means-based speckle filtering for ultrasound images’, *IEEE Trans. on Image Process.* **18**(10), 2221–2229.
- Coupé, P., Hellier, P., Prima, S., Kervrann, C. & Barillot, C. (2008), ‘3d wavelet sub-bands mixing for image denoising’, *Int. J. Biomedical Imaging* **8**(3), 1–11.
- Coupé, P., Yger, P., Prima, S., Hellier, P., Kervrann, C. & Barillot, C. (2008), ‘An optimized blockwise non-local means denoising filter for 3d magnetic resonance images’, *IEEE Trans. on Medical Imaging* **27**(4), 325–441.
- Coupé, P., Hellier, P., Kervrann, C. & Barillot, C. (2008), Bayesian non-local means-based speckle filtering, *in* ‘Proc. IEEE Int. Symp. on Biomedical Imaging : from nano to macro (ISBI’08)’, Paris, France, pp. 291–294.

-
- Kervrann, C. (1998), ‘Learning probabilistic deformation models from image sequences’, *Signal Process.* (71), 155–171.
- Kervrann, C. (2001), ‘Bayesian image segmentation through level lines selection’, *Image Anal. and Stereology* **20**(1), 251–256.
- Kervrann, C. (2004), An adaptive window approach for image smoothing and structures preserving, in ‘Proc. Eur. Conf. Comput. Vision (ECCV’04) (3)’, Prague, Czech Republic, pp. 132–144.
- Kervrann, C. & Boulanger, J. (2006a), ‘Optimal spatial adaptation for patch-based image denoising’, *IEEE Trans. on Image Process.* **15**(10), 2866–2878.
- Kervrann, C. & Boulanger, J. (2006b), Unsupervised patch-based image regularization and representation, in ‘Proc. Eur. Conf. Comput. Vision (ECCV’06)’, Graz, Austria, pp. 555–567.
- Kervrann, C. & Boulanger, J. (2008), ‘Local adaptivity to variable smoothness for exemplar-based image denoising and representation’, *Int. J. Comput. Vision* **79**(1), 45–69.
- Kervrann, C., Boulanger, J. & Coupé, P. (2007), Bayesian non-local means filter, image redundancy and adaptive dictionaries for noise removal, in ‘Proc. Conf. on Scale-Space and Variational Meth. (SSVM’07)’, Ischia, Italy, pp. 520–532.
- Kervrann, C., Boulanger, J., Pécot, T., Pérez, P. & J. Salameró, J. (2009), ‘Neighborhood-wise multiscale decision fusion for redundancy detection in image pairs’, Submitted to SIAM J. Multiscale Modeling & Simul.
- Kervrann, C., Boulanger, J., Pécot, T. & Pérez, P. (2009), Discriminant random field and patch-based redundancy analysis for image change detection, in ‘Proc. IEEE Int. Workshop on Mach. Learning for Signal Process. (MLSP’09)’, Grenoble, France, pp. 1-6.
- Kervrann, C. & Heitz, F. (1994), ‘Segmentation non supervisée des images naturelles texturées : une approche statistique’, *Traitement du Signal* **11**(1), 31–41.
- Kervrann, C. & Heitz, F. (1995), ‘A Markov random field model-based approach to unsupervised texture segmentation using local and global spatial statistics’, *IEEE Trans. on Image Process.* **4**(6), 856–862.
- Kervrann, C. & Heitz, F. (1998), ‘A hierarchical Markov modeling approach for the segmentation and tracking of deformable shapes’, *Graphical Models and Image Process.* **60**(3), 173–195.
- Kervrann, C. & Heitz, F. (1999), ‘Statistical deformable model-based segmentation of image motion’, *IEEE Trans. on Image Process.* **8**(4), 583–588.
- Kervrann, C., Hoebeke, M. & Trubuil, A. (1999), A level line selection approach for object boundary estimation, in ‘Proc. IEEE Int. Conf. Comput. Vision (ICCV’99)’, Kerkyra, Greece, pp. 963–968.
- Kervrann, C., Hoebeke, M. & Trubuil, A. (2000), Level lines as global minimizers of energy functionals in image segmentation, in ‘Proc. Eur. Conf. Comput. Vision (ECCV’00)’, Dublin, Ireland, pp. 241–256.
- Kervrann, C., Hoebeke, M. & Trubuil, A. (2002), ‘Isophotes selection and reaction-diffusion model for object boundaries estimation’, *Int. J. Comput. Vision* **50**(1), 63–94.

- Kervrann, C., Legland, D. & Pardini, L. (2004), ‘Robust incremental compensation of the light attenuation with depth in 3d fluorescence microscopy’, *J. Microscopy* **214**(3), 297–314.
- Kervrann, C. & Trubuil, A. (2002), ‘Optimal level curves and minimizers of cost functionals in image segmentation’, *J. Math. Imaging and Vision* **17**(2), 153–174.
- Kervrann, C. & Trubuil, A. (2004), An adaptive window approach for poisson noise reduction and structure preserving in confocal microscopy, *in* ‘Proc. IEEE Int. Symp. on Biomedical Imaging : from nano to macro (ISBI’04)’, Arlington, USA, pp. 788-791.
- Pécot, T., Chessel, A., Bardin, S., Salamero, J., Bouthemy, P. & Kervrann, C. (2009), Conditional random fields for object and background estimation in fluorescence video-microscopy, *in* ‘Proc. IEEE Int. Symp. on Biomedical Imaging : from nano to macro (ISBI’09)’, Boston, USA, pp. 734–737.
- Pécot, T., Boulanger, J., Kervrann, C. & Bouthemy, P. (2007), Network tomography for trafficking simulation and analysis in fluorescence microscopy imaging, *in* ‘Proc. IEEE Int. Symp. on Biomedical Imaging : from nano to macro (ISBI’07)’, Washington, USA, pp. 268–271.
- Pécot, T., Kervrann, C., Bardin, S., Goud, B., & Salamero, J. (2008), Patch-based Markov models for event detection in fluorescence bioimaging, *in* ‘Proc. Int. Conf. on Medical Image Computing and Computer Assisted Intervention (MICCAI’08)’, New York City, USA, pp. 95-103.
- Pécot, T., Kervrann, C. & Bouthemy, P. (2008a), Minimal paths and probabilistic models for origin-destination traffic estimation in live cell imaging, *in* ‘Proc. IEEE Int. Symp. on Biomedical Imaging : from nano to macro (ISBI’08)’, Paris, France, pp. 843–846.
- Pécot, T., Kervrann, C. & Bouthemy, P. (2008b), Network tomography-based tracking for intracellular traffic analysis in fluorescence microscopy imaging, *in* ‘Proc. Int. Conf. on Bio-inspired Systems and Signal Process. (BIOSIGNALS’08) (1)’, Funchal, Portugal, pp. 154–161.
- Racine, V., Hertzog, A., Jouaneau, J., Salamero, J., Kervrann, C. & Sibarita, J.-B. (2006), Multiple target tracking of 3D fluorescent objects based on simulated annealing, *in* ‘Proc. IEEE Int. Symp. on Biomedical Imaging : from nano to macro (ISBI’06)’, Washington, USA, pp. 1020–1023.