



HAL
open science

Outils d'analyse d'images et recalage d'individus pour l'étude de la morphogenèse animale et végétale

Gaël Michelin

► **To cite this version:**

Gaël Michelin. Outils d'analyse d'images et recalage d'individus pour l'étude de la morphogenèse animale et végétale. Autre. COMUE Université Côte d'Azur (2015 - 2019), 2016. Français. NNT : 2016AZUR4088 . tel-01451608

HAL Id: tel-01451608

<https://theses.hal.science/tel-01451608>

Submitted on 1 Feb 2017

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.

UNIVERSITÉ CÔTE D'AZUR
ÉCOLE DOCTORALE STIC
SCIENCES ET TECHNOLOGIES DE L'INFORMATION
ET DE LA COMMUNICATION

THÈSE

pour obtenir le titre de

Docteur en Sciences

de l'Université Côte d'Azur

Mention : AUTOMATIQUE ET TRAITEMENT DU SIGNAL

Présentée et soutenue par

Gaël MICHELIN

Outils d'analyse d'images et recalage d'individus pour l'étude de la morphogenèse animale et végétale

Thèse dirigée par Grégoire MALANDAIN

préparée à Inria Sophia Antipolis, Projet MORPHEME

soutenue le 28 octobre 2016

Jury :

<i>Président :</i>	Jan TRAAS	- ENS Lyon
<i>Rapporteurs :</i>	Philippe ANDREY	- Inra
	Patrick BOUTHEMY	- Inria Rennes
<i>Directeur :</i>	Grégoire MALANDAIN	- Inria Sophia Antipolis
<i>Examineurs :</i>	Christophe GODIN	- Inria Montpellier
	Alain TRUBUIL	- Inra
<i>Invité :</i>	Patrick LEMAIRE	- CNRS

Remerciements



J'adresse mes premiers remerciements à Grégoire Malandain. Merci pour la confiance que tu m'as accordée au cours de ces trois années. Trois ans, que le temps passe... Encore un peu de patience, d'ici peu tu auras dépassé l'âge de mes parents (sic)! Et qui sait, lorsque ce sera le cas, tu auras peut-être une grosse moto...

Parmi les personnes avec qui j'ai eu le plaisir et la chance de collaborer au cours de ces années, je remercie très chaleureusement Christophe Godin. J'ai été constamment impressionné par ton implication, toujours pertinente et enrichissante, dans mon travail, mais aussi sur tous les autres fronts du projet Morphogenetics. Je remercie également Jan Traas, qui a toujours su créer une ambiance de travail très dynamique et stimulante au sein du projet Morphogenetics. Enfin, je remercie Patrick Lemaire, non seulement au titre de notre collaboration passée et à venir, mais aussi au titre de son engagement précieux dans la lutte pour protéger la recherche scientifique en France. Merci à vous trois pour votre amitié.

Je tiens à remercier les membres restants de mon jury de thèse, Philippe Andrey, Patrick Bouthemey et Alain Trubuil. Merci pour votre effort de relecture, vos conseils, votre écoute et même parfois vos encouragements fort appréciables en particulier à l'approche de l'échance « finale ».

Bien sûr, je remercie également l'ensemble des collègues chercheurs, ingénieurs, stagiaires, doctorants avec lesquels j'ai été amené à travailler, avec une attention toute particulière pour Léo Guignard et Yassin Refahi. Je n'oublie cependant pas les échanges que j'ai pu avoir avec Ulla-Maj Fiuza, Emmanuel Faure, Pradeep Daz, Annamaria Kiss, Henrik Jönsson, Raymond Wightman, ainsi que l'ensemble des membres de l'équipe Virtual Plants qui a toujours bien su m'accueillir lors de mes passages à Montpellier.

Mon expérience au sein de l'équipe Morpheme fût en tous points formidable, qu'il s'agisse des expéditions lointaines en conférence ou des fameux « Morpheme Ski Days » en passant par les pique-niques d'équipe, les repas, les discussions très

intellectuelles, les activités bricolage, les parties de pétanque et diverses expérimentations « pour la science » dont je fus parfois le malheureux cobaye (voir figure 1).

Je remercie en premier lieu l'artisan, le pilier de cette équipe, Xavier. Je remercie également les valeureux Grégoire, Laure et Eric côté Sophipolitain, ainsi que Caroline (pas toi Maman, une autre) et Florence aux ailes (Valrose). Merci également au staff de l'équipe, recomposé en cours de route mais toujours au top : Frédéric, Janette, Micheline, Nadia et Christine. Bien sûr, je pense aussi aux courageux « anciens » de l'équipe : Nektaria, Anirudh, Mériadec, Huei-Fang, Alejandro, Rita, Simone, Simon, Raphaël, Nicoco, la Chinese team (Yaqun, Sen, Wei), Tarun, Marine Ma Divine et Manon :3. Je pense également aux amateurs de troisième mi-temps : Djampa, Agus (rrrRRRrrr), Lolita, Emma, Anca. Enfin, je pense à mes deux coéquipiers de luxe Alexis/ALPHA et Manu/PICA. Puissent les valeurs profondément rugbystiques instaurées dans l'équipe Morpheme perdurer.

Merci aussi à l'équipe des grimpeurs/joueurs : Marine (cumul d'activités avec Morpheme), Ludo, Mélaine.

Enfin, on n'a pas la même équipe, mais on a les mêmes valeurs : Ophélie, Benjamin, Emilien, Jonathan, mais aussi Cédric (auf Wiedersehen), Froso (shame on you) et Yasmine.

Je pense aussi à l'équipe Vortex de l'IRIT à Toulouse, qu'il s'agisse de Vincent, de Jean-Denis, de Sylvie, ou encore de Géraldine. C'est à travers votre équipe que j'ai découvert le « monde » académique, vous m'avez beaucoup apporté lors de ma formation et je vous en suis aujourd'hui très reconnaissant.

Au delà de mon environnement de travail, je remercie l'ensemble de mes proches, parmi lesquels l'équipe des baroudeurs avec qui j'ai eu l'opportunité de découvrir de magnifiques régions d'Europe et d'Amérique du Sud durant mes années de doctorat en attendant de nouvelles destinations : Clémence, Roland, Mathieu, Stéphane, Lesley, Nico, Sébastien, Laurence. Merci également à Yuravin et Vincent pour votre venue à ma soutenance. Enfin, un simple mot adressé à Yoann : wololo !

Je clos cette liste de remerciements avec une pensée très affectueuse pour ma famille, qui m'a toujours soutenu et que je chéris par dessus tout. Merci à mes parents, à mon frère Nicolas, ma sœur Fanny, mes oncles et tantes, mes grand-mères Mémé et Mamie Dédée que j'embrasse très fort. Le mot de la fin sera pour mes cousins. À Vincent, Martin, Dimitri, Benjamin : merci beaucoup.



FIGURE 1 – Poison.

Résumé

En biologie développementale, l'étude d'organismes modèles vise à comprendre les mécanismes génétiques responsables de la morphogenèse chez le vivant. La microscopie confocale à fluorescence permet aujourd'hui d'observer *in vivo* à l'échelle de la cellule et avec une haute fréquence temporelle le développement d'organismes. Les séquences d'images 3D+ t ainsi obtenues nécessitent d'avoir des outils de traitement d'images adaptés.

Dans cette thèse, nous construisons des outils dédiés à l'étude du développement de deux organismes, l'embryon de l'ascidie *Phallusia mammillata* et le bouton floral d'*Arabidopsis thaliana*.

Nous développons d'abord une méthode de comparaison de segmentations adaptée aux images de tissus épithéliaux d'organismes en développement. Nous nous appuyons sur cet outil pour valider notre seconde contribution qui porte sur la mise en place d'un outil de détection et de reconstruction de membranes cellulaires conçu afin de procéder à la segmentation de cellules dans les images d'ascidies et d'arabidopsis.

Nous utilisons ensuite l'outil de segmentation de membranes précédemment introduit pour construire une stratégie de recalage spatial inter-individus appliquée aux embryons d'ascidies. Enfin, nous élaborons une stratégie de recalage spatio-temporel inter-individus appliquée à des séquences d'images 3D de méristèmes floraux.

Les outils développés ouvrent la perspective de construire des modèles digitaux d'organismes en développement sur lesquels des populations d'individus peuvent être projetées et ainsi de conduire des études comparatives quantifiées sur ces populations.

Mots clés : Biologie développementale, microscopie à fluorescence, segmentation, recalage, ascidie, arabidopsis

Abstract

In developmental biology, the study of model organisms aims for the understanding of genetic mechanisms responsible of morphogenesis. Today, fluorescent confocal microscopy is a means for *in vivo* imaging of developing organisms at cell level with a high spatio-temporal resolution. To handle such $3D+t$ image sequences, adapted computer-assisted methods are highly desirable.

In this thesis, we build dedicated tools for the study of two developing organisms, the ascidian *Phallusia mammillata*'s embryo and the *Arabidopsis thaliana*'s floral meristem.

We first develop a method for segmentation comparison adapted to developing organism epithelial tissue images. This tool is then used to validate our second contribution that is about the development of a cell membranes detection and reconstruction tool for cell shape segmentation process applied to ascidian and arabidopsis images.

We then use the previously introduced membrane detection tool to build an inter-individual spatial registration strategy applied to ascidian embryo images. Finally, we develop an inter-individual spatio-temporal registration strategy applied to 3D image sequences of arabidopsis floral meristems.

The developed tools offers the means to build 4D digital templates of developing organisms on which every individual can be projected, opening the avenue to the statistical analysis of populations.

Keywords : Developmental biology, fluorescence microscopy, segmentation, registration, ascidian, arabidopsis

Table des matières

1	Avant-propos	1
2	Introduction générale	3
2.1	Biologie développementale	3
2.1.1	Définition	3
2.1.2	Des organismes modèles	4
2.2	Imagerie	8
2.2.1	Microscopie par fluorescence	8
2.2.2	Marquage des structures	10
2.3	Analyse d'images	12
2.3.1	Pré-traitements	12
2.3.2	Détection et segmentation des noyaux et des cellules	13
2.3.3	Lignées cellulaires	15
2.3.4	Quantification de l'expression des gènes	18
2.4	Plan de la thèse	19
3	Comparaison de segmentations	21
3.1	Motivations	21
3.2	Cas de figure	24
3.3	Mesure de recouvrement de régions	25
3.3.1	Taux d'inclusion de l'érodé	25
3.3.2	Relation avec des indices symétriques	26
3.4	Matrices d'intersections de régions labellisées	27
3.5	Association des labels	28
3.6	Expériences	29
3.6.1	Données de test	29
3.6.2	Résultats	31
3.7	Conclusion	36
4	Segmentation de membranes cellulaires	37
4.1	Introduction	37
4.2	Fonction de planéité	39
4.2.1	Introduction	39
4.2.2	Modèle planaire	43
4.2.3	Fonction de réponse	43
4.3	Binarisation des membranes	45
4.3.1	Extraction des plans centraux	45
4.3.2	Inhomogénéité de l'intensité des membranes	47
4.3.3	Fonction de seuillage adaptatif	47
4.3.4	Histogrammes axiaux	48

4.3.5	Seuils axiaux	49
4.4	Reconstruction d'une image d'intensité par vote de tenseurs	50
4.4.1	Introduction	50
4.4.2	Les votes de tenseurs	51
4.5	Segmentation des cellules	57
4.5.1	Détection de graines	58
4.5.2	Ligne de partage des eaux	58
4.6	Expériences et résultats	58
4.6.1	Données	58
4.6.2	Évaluation de la binarisation des membranes	59
4.6.3	Évaluation du pré-traitement	61
4.6.4	Saillance membranaire	64
4.7	Conclusions et perspectives	65
5	Recalage d'embryons d'ascidies	69
5.1	Introduction	69
5.2	Recalage d'embryons segmentés : notre approche	70
5.3	Détection de la symétrie bilatérale de l'embryon	74
5.3.1	Extraction des axes de symétrie bilatérale candidats	76
5.3.2	Initialisation des plans de symétrie candidats	76
5.3.3	Sélection du meilleur plan candidat	78
5.3.4	Ajustement de l'interface de symétrie bilatérale par un plan	78
5.4	Recalage d'embryons d'ascidies	79
5.4.1	Initialisation des appariements	81
5.4.2	Itérations : choix de la fonction de coût	82
5.4.3	Solution globale	83
5.5	Expériences et résultats	83
5.5.1	Données	83
5.5.2	Extraction du plan de symétrie bilatérale	85
5.5.3	Recalage d'embryons	88
5.6	Conclusions et perspectives	91
6	Recalage spatio-temporel de séquences 3D+t de méristèmes floraux	95
6.1	Introduction	96
6.2	Augmentation de la résolution temporelle d'une séquence d'images 3D de méristème en développement	99
6.2.1	Compensation du déplacement rigide de la plante au fil des acquisitions	99
6.2.2	Recalage non linéaire entre paires d'images	100
6.2.3	Construction d'une séquence d'images interpolées	100
6.3	Recalage spatial de méristèmes floraux par une approche géométrique	103
6.3.1	Extraction de primitives géométriques	104
6.3.2	Recalage par la méthode du plus proche point itéré	104

6.3.3	Complexité algorithmique et choix d'implémentation	107
6.4	Recalage temporel d'une image sur une séquence	107
6.4.1	Extraction des primitives d'une image interpolée	108
6.4.2	Sélection de la meilleure correspondance temporelle	108
6.4.3	Critère de similarité	109
6.5	Stratégies de recalage de séquences de MF	113
6.5.1	Notations	113
6.5.2	Recalage image sur image	114
6.5.3	Recalage image sur séquence	115
6.5.4	Recalage séquence sur séquence	118
6.6	Expériences et résultats	119
6.6.1	Données	119
6.6.2	Interpolation d'images	121
6.6.3	Recalage spatial d'images de MF	123
6.6.4	Stratégies de recalage	130
6.7	Conclusions et perspectives	131
7	Conclusion générale	139
7.1	Contributions	139
7.2	Discussions et perspectives	141
7.2.1	Segmentation d'images	141
7.2.2	Recalage d'organismes en développement	141
7.2.3	Mise à disposition des outils développés	144
A	Résultats complets de détection des cellules	145
B	Recalage multimodal d'images d'ascidies	149
B.1	Description des données	149
B.2	Estimation de l'affaissement de l'embryon test	149
B.3	Détection de la symétrie pour l'image test	150
B.4	Recalage multimodal	150
B.5	Conclusion	155
C	Interpolation multi-forme	157
C.1	Notations	157
C.2	Principe	157
C.3	Lien avec l'interpolation basée sur la forme (<i>Shape-based interpolation</i>)	158
C.4	Détails d'implémentation	159
D	Épaisseur de la couche L1 du MF d'Arabidopsis	161
	Bibliographie	165

Avant-propos

Mon travail a été financé par une bourse CORDI-S dans le cadre de l’Inria Project Lab (IPL) Morphogenetics. Ce projet de recherche a pour but d’éclairer le mécanisme de régulation génétique lors du développement des organismes, en particulier les organes floraux. Les recherches menées par ce projet sont transversales. Elles recouvrent des aspects génétiques, moléculaires, géométriques ou encore mécaniques au service de modèles computationnels qui permettent de simuler la croissance des cellules et des tissus au sein de la fleur.

L’IPL Morphogenetics regroupe les équipes-projets Virtual Plants, Imagine et Morpheme d’Inria ainsi que les équipes Méristème et Biophysique & Développement de l’UMR Reproduction et Développement des Plantes (ENS Lyon, Inra, CNRS) et l’équipe Régulateurs du développement de la plante (UMR Physiologie Cellulaire Végétale CEA, Inra, CNRS, Université Grenoble Alpes). Une collaboration étroite entre ces différentes équipes décuple notablement l’efficacité du travail de recherche de chacune des équipes du projet. Le partage d’information, facilité par des réunions de travail à intervalles réguliers à l’état d’esprit constructif, permet d’orchestrer de manière réfléchie et conjointe le travail de l’ensemble des ressources humaines dont le projet dispose.

Parmi les axes de recherche du projet Morphogenetics, mes travaux visent à répondre au besoin de construire des modèles numériques spatio-temporels d’organismes en développement et de fournir aux biologistes des outils de mesures quantitatives sur des populations. La finalité de cet axe de recherche est de construire des organismes *in silico*, c’est-à-dire des organismes complètement virtuels (numériques) qui permettent de simuler et de prédire le comportement réel d’un organisme en développement selon son génome et son environnement.

Mon sujet de thèse s’est focalisé sur l’étude de deux organismes que sont l’*Arabidopsis* (plus précisément son méristème apical) et l’ascidie (animal marin) lors de son embryogenèse. Les techniques d’imagerie récemment mises au point permettent aujourd’hui d’observer le développement de ces organismes à l’échelle cellulaire par le biais de séquences temporelles d’images 3D. Compte-tenu de l’explosion des quantités d’images à traiter, il devient indispensable d’automatiser les tâches de traitement et d’analyse de ces données via des outils informatiques. Les contributions que j’apporte dans ce mémoire portent sur des aspects de traitement d’image (segmentation ou identification de cellules, recalage ou alignement d’individus) et d’analyse (comparaison de segmentations). Les méthodes développées au cours de mon travail sont testées sur des données réelles, ce qui permet de mesurer les points forts et faiblesses de ces outils dans un contexte pratique.

Introduction générale

Sommaire

2.1	Biologie développementale	3
2.1.1	Définition	3
2.1.2	Des organismes modèles	4
2.2	Imagerie	8
2.2.1	Microscopie par fluorescence	8
2.2.2	Marquage des structures	10
2.3	Analyse d'images	12
2.3.1	Pré-traitements	12
2.3.2	Détection et segmentation des noyaux et des cellules	13
2.3.3	Lignées cellulaires	15
2.3.4	Quantification de l'expression des gènes	18
2.4	Plan de la thèse	19

2.1 Biologie développementale

2.1.1 Définition

La biologie développementale est un thème des sciences du vivant qui regroupe l'ensemble des études visant à répondre à la question « Comment une simple cellule-œuf peut-elle évoluer et former un organisme complexe? ». Dans ce but, les recherches sur l'embryogenèse, c'est-à-dire les processus de croissance et de développement des organismes multi-cellulaires, cherchent à éclaircir le lien entre la génétique et la variation phénotypique au niveau des individus, des populations et des espèces.

L'apparition de la microscopie et son utilisation en biologie marquent un tournant définitif dans l'approche de cette science, avec la découverte des cellules attribuée à Robert Hooke au cours du XVII^{ème} siècle. Aujourd'hui, l'imagerie s'impose comme le principal moyen d'investigation pour la compréhension de l'embryogenèse [Megason 2007], dont les processus principalement étudiés sont :

- la différenciation cellulaire, processus permettant aux cellules de se spécialiser bien qu'ayant le même patrimoine génétique ;
- le contrôle génétique de la croissance cellulaire ;
- la morphogenèse, processus de développement des structures d'un organisme au cours de son embryogenèse.

2.1.2 Des organismes modèles

2.1.2.1 Définition

En biologie, un organisme modèle est une espèce étudiée de manière approfondie afin de comprendre un éventail de phénomènes biologiques, dans l'espoir que les résultats obtenus par l'expérimentation sur ce modèle soient applicables à d'autres organismes potentiellement plus complexes que le modèle [Bolker 1995, Ankeny 2011].

Le recours à des organismes modèles est justifié en raison, d'une part de l'universalité de la molécule d'ADN, d'autre part de l'unité cellulaire du vivant et de son unicité structurale et fonctionnelle. Aussi, certaines espèces ont été retenues car elles partagent plusieurs caractéristiques intéressantes du point de vue génétique, expérimental et économique :

- génome bien caractérisé, souvent de faible taille, outils de biologie moléculaire disponibles pour son séquençage et sa manipulation à l'aide de diverses techniques ;
- temps de génération relativement faible, descendance nombreuse ;
- facilité d'élevage ou de culture, faible encombrement.

Il convient de distinguer l'animal du végétal en biologie développementale. Chez l'animal, l'organogenèse s'établit au cours de l'embryogenèse et dès lors, le développement post-embryonnaire de l'organisme poursuit le plan précédemment établi, sans formation de nouvel organe excepté chez les animaux possédant un stade de larve dans leur développement. À l'inverse, la plante maintient tout au long de son développement sa capacité de créer de nouveaux organes à l'aide de quelques milliers de cellules souches formant des structures microscopiques, appelées méristèmes. Cette faculté d'organogenèse continue au cours de leur vie distingue les plantes des animaux.

2.1.2.2 L'ascidie comme organisme modèle chez l'animal

« ... at an extremely remote period a group of animals existed, resembling in many respects the larvae of our present Ascidians, which diverged into two great branches—the one retrograding in development and producing the present class of Ascidians, the other rising to the crown and summit of the animal kingdom by giving birth to the Vertebrata. »

[Darwin 1888]

Les ascidies constituent une classe d'animaux marins invertébrés appartenant au sous-embranchement des tuniciers qui sont voisins des vertébrés. On trouve ces animaux dans tous les océans du monde, et en particulier dans les eaux peu profondes des littoraux où elles se fixent sur les roches, les fonds littoraux ou encore les coques de bateaux. À l'âge adulte, les ascidies sont des animaux benthiques filtreurs qui se nourrissent des nutriments contenus dans l'eau qui les traverse.

Le développement des embryons d'ascidies est rapide et stéréotypé, avec des lignées cellulaires invariantes et une symétrie bilatérale du stade d'œuf jusqu'au



FIGURE 2.1 – Dessin de Saville Kent (1860) représentant des ascidies. 1 Ascidie blanche ou *Phallusia mammillata* ; 2 *Pyura microcosmus* ; 3 Cione ou *Ciona intestinalis* ; 4 Ascidie rouge ou *Halocynthia papillosa* ; 5 Botrylle étoilé ou *Botryllus schlosseri* ou *Polycyclus renieri* ; 6 *Botrylloides rubrum*. [Waterwereld 2002]

stade de têtard (ou larve) [Lemaire 2009]. Toutes les espèces d'ascidies étudiées présentent un développement embryonnaire quasiment identique. Ce n'est qu'à l'âge adulte que des métamorphoses complexes résultent en une grande variété de formes selon les espèces d'ascidies (figure 2.1).

Le développement embryonnaire se décompose en plusieurs grandes périodes (segmentation, prégastrulation, gastrulation, neurulation, organogenèse). Jusqu'au stade de gastrulation, au cours duquel on assiste à l'invagination de l'embryon et à la mise en place du tube digestif, les stades de développement d'un embryon d'ascidie sont déterminés par le nombre de cellules le composant. Puis ils sont définis par les événements majeurs de leur développement que sont la gastrulation, la neurulation, la croissance de la queue, jusqu'à ce que l'embryon atteigne le stade de larve éclore (figure 2.2).

L'ascidie est un un modèle privilégié pour le déchiffrage et l'analyse des réseaux complexes de régulation des gènes du développement [Imai 2006] en raison de sa simplicité, son invariance, son faible nombre de cellules et la détermination rapide de leur destinée. Les génomes de plusieurs espèces d'ascidies ont été séquencés et sont accessibles via le portail ANISEED, permettant de caractériser des différences génomiques intra- et inter-espèces [Dehal 2002, Kim 2007] et d'analyser leurs répercussions sur les systèmes de régulation [Stolfi 2014]. La taille et la transparence de l'ascidie (en particulier *Phallusia mammillata*) rendent possible une imagerie à haute résolution spatiale et temporelle permettant d'observer à l'échelle de la cellule la croissance de l'embryon, avec la possibilité de segmenter et d'extraire les lignées cellulaires [Guignard 2015].

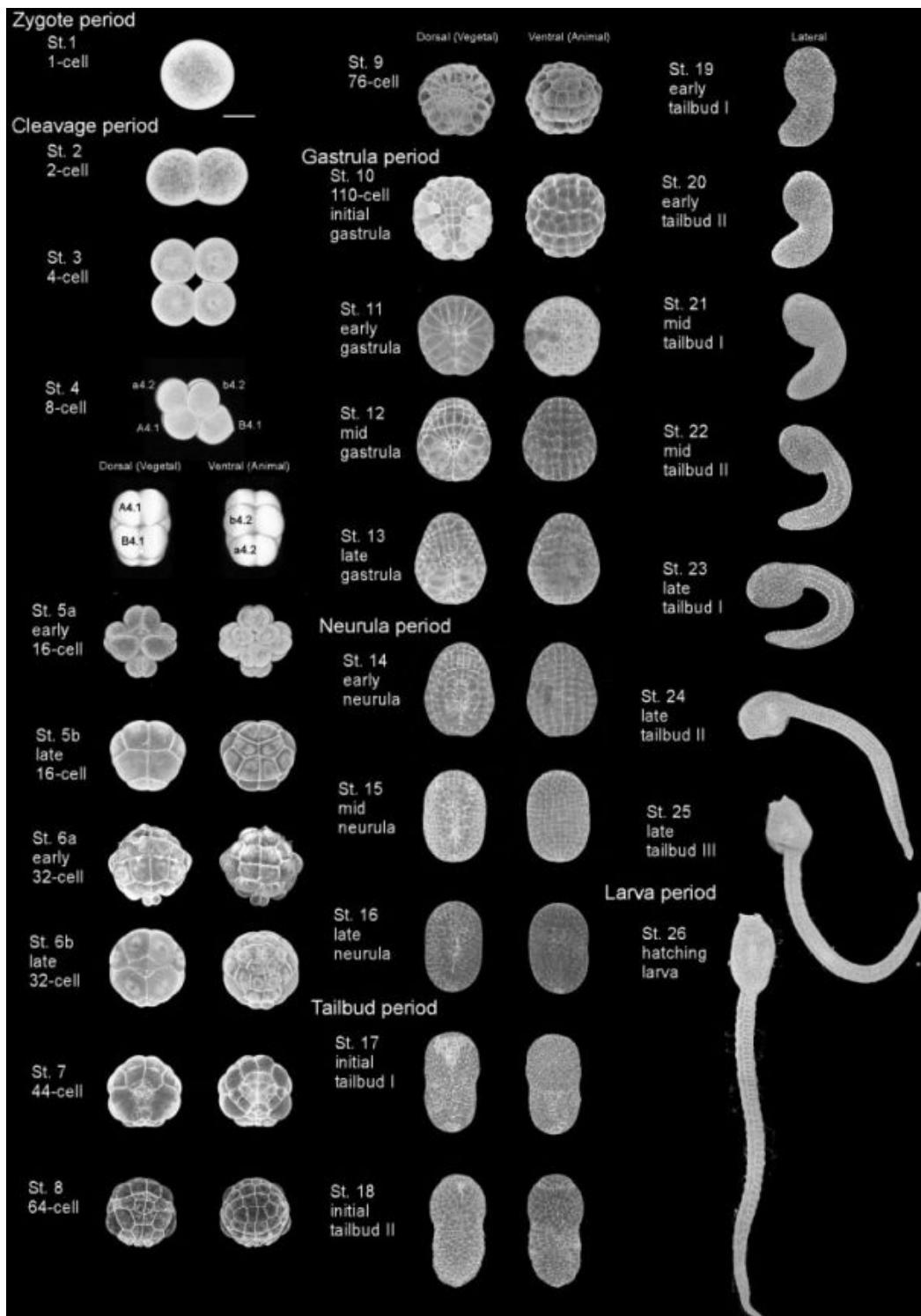


FIGURE 2.2 – Stades de développement de l'ascidie *Ciona intestinalis*, du stade d'oeuf au stade de larve [Hotta 2007].



FIGURE 2.3 – Plant d'Arabidopsis (gauche) et observation au microscope du méristème apical et de boutons floraux (droite).

2.1.2.3 Arabidopsis comme organisme modèle chez la plante

Dans le règne végétal, *Arabidopsis thaliana* appartient à la famille des Brassicacées (sous-groupe des dicotylédones). Largement répandue à travers le monde et qualifiée de « mauvaise herbe » par tout un chacun, il s'agit d'une plante annuelle mesurant 20 à 25 cm de hauteur à l'âge adulte et ornée de fleurs blanches d'un diamètre de 3 mm (figure 2.3).

L'histoire commune entre *Arabidopsis thaliana* et les biologistes date du début du XX^{ème} siècle. En 1943, Friedrich Laibach suggère pour la première fois l'utilisation de l'arabidopsis comme un organisme modèle [Laibach 1943]. Finalement, c'est dans les années 1990 que l'arabidopsis s'impose comme telle. Aujourd'hui, elle revêt une importance comparable à celle des drosophiles ou des souris en biologie animale, avec un nombre considérable de publications scientifiques faisant mention du terme "Arabidopsis" dans leur titre, leur résumé ou leurs mots-clés (plus de 50 000 publications au cours de ces 50 dernières années) [Provar 2016]. Aujourd'hui, l'arabidopsis est largement utilisée pour la recherche génétique et développementale. Il s'agit de la première plante dont le génome a été complètement séquencé. L'intérêt pour l'arabidopsis s'explique par :

- la taille de son génome, parmi les plus petits du monde végétal ;
- son cycle de vie très court (environ 2 mois) ;
- sa petite taille idéale pour sa culture en laboratoire ;
- l'absence d'enjeu agronomique, ce qui la préserve de toute tentative d'appropriation industrielle et facilite le partage d'information entre laboratoires.

2.2 Imagerie

En biologie développementale, l'imagerie se décline en deux types d'observations. Les observations de type morphologique permettent d'accéder à la forme des tissus cellulaires. Les observations de type fonctionnel ont pour but d'étudier la physiologie du tissu (mécanique, physique et biochimie d'un organisme). Dans ce manuscrit, nous considérerons exclusivement des observations de type morphologique sur des tissus vivants. Dans la majorité des cas, les structures observées sont les noyaux et les membranes des cellules. Imager les noyaux permet de réaliser une détection et un suivi du déplacement des cellules, mais n'offre pas une information précise sur la forme des cellules, alors que l'imagerie des membranes permet d'accéder à la morphologie des cellules. Dans de nombreux cas, l'imagerie en $3D+t$ de ces structures est une tâche complexe : tissus opaques, absorption et dispersion de la lumière, sur-exposition de l'organisme à la lumière entraînant une perturbation de son développement... Les techniques d'imagerie et les protocoles d'acquisition sont constamment revus afin d'être moins invasifs et plus précis. Lorsqu'il est possible de marquer les structures avec des fluorophores, la microscopie par fluorescence s'impose comme un moyen efficace d'imagerie dans le cadre d'études sur la morphogenèse par le biais de la microscopie confocale et la microscopie bi-photon [Keller 2013].

Les avancées en imagerie cellulaire *in vivo* non-invasive rendant possible la quantification des activités cellulaires sont décrites dans [Liu 2015]. La suite de cette section présente brièvement les techniques d'imagerie et de marquage ayant permis d'acquérir les données sur lesquelles je me suis appuyé au cours de ma thèse.

2.2.1 Microscopie par fluorescence

En microscopie optique classique, l'observation se fait par réflexion ou absorption de la lumière émise par une source non ponctuelle (rétro-éclairage du tissu). Cette technique comporte des limitations, comme l'épaisseur du tissu observé qui doit être faible ou la présence d'objets parasites à proximité du plan focal. La microscopie par fluorescence lève ces restrictions et permet d'observer des organismes naturellement fluorescents ou rendus fluorescents par un marqueur pouvant être génétique, comme la Green Fluorescent Protein (GFP) [Chalfie 1994], dont la découverte fut récompensée par le Prix Nobel de Chimie en 2008 [Nobelprize.org 2014].

Les différents microscopes à fluorescence émettent de la lumière à une longueur d'onde donnée qui excite les fluorophores présents dans le tissu observé. À leur tour, les fluorophores émettent de la lumière à une longueur d'onde plus grande que celle de la source et un capteur collecte la lumière émise par l'ensemble du tissu excité. On distingue plusieurs types de microscopes à fluorescence qui se distinguent principalement par leur méthode de balayage de l'échantillon et par leur source d'excitation.

2.2.1.1 Microscopie confocale

Les microscopes à fluorescence confocaux utilisent une source de lumière qui excite l'échantillon sur des sections optiques de faible épaisseur (environ 400 nm) au niveau du plan focal de l'objectif. En déplaçant la position du plan focal, il est alors possible d'imager l'échantillon à différents niveaux de profondeur et donc de construire une image tri-dimensionnelle.

Microscopie confocale à balayage laser. En microscopie confocale à balayage laser, l'échantillon est balayé point par point par un faisceau de laser (figure 2.4a). Un sténopé placé devant le dispositif de détection de lumière bloque l'ensemble des photons émis hors du plan focal, de telle sorte que seuls les photons provenant des molécules excitées du plan focal participent à la formation de l'image. Bien que permettant un balayage précis de l'échantillon, cette méthode d'imagerie comporte plusieurs défauts. Principalement, la résolution temporelle de l'acquisition peut être relativement faible lors d'acquisitions nécessitant le balayage de nombreux points. De plus, le balayage de chaque point de l'échantillon par le faisceau laser nécessite d'illuminer l'épaisseur entière de l'échantillon, ce qui provoque une sur-exposition à la lumière de l'organisme et donc à la fois une photo-toxicité accrue pouvant affecter son développement et un phénomène de photoblanchiment (perte de fluorescence des molécules).

Microscopie à feuille de lumière (SPIM). En microscopie confocale, parmi plusieurs alternatives au microscope à balayage laser améliorant certains aspects de l'acquisition d'image, nous citerons en particulier la méthode de microscopie à feuille de lumière qui permet d'obtenir une excellente résolution temporelle et de diminuer significativement la photo-toxicité de l'imagerie de l'échantillon [Huisken 2004]. Cette technique consiste à illuminer et imager l'ensemble du plan focal de l'échantillon afin de construire coupe par coupe l'image en 3D en déplaçant le plan focal (figure 2.4b). À l'inverse de la méthode par balayage laser, l'ensemble des molécules ne sont illuminés qu'une fois (lorsqu'elles sont situées dans le plan focal), ce qui permet une diminution du photoblanchiment et de la photo-toxicité induits par l'imagerie. Le temps nécessaire à l'acquisition est également considérablement amélioré, puisqu'il ne dépend plus que du nombre de coupes imagées (au lieu du nombre total de points parcourus en balayage laser). Cependant, le gain en vitesse et photoblanchiment est contre-balancé par une perte en focus. Si le tissu observé n'est pas suffisamment transparent, son imagerie nécessitera un laser plus puissant qui induira une photo-toxicité potentiellement importante. La microscopie à feuille de lumière est aujourd'hui largement répandue en biologie cellulaire et en particulier dans le contexte du développement animal [Keller 2010, Krzic 2012].

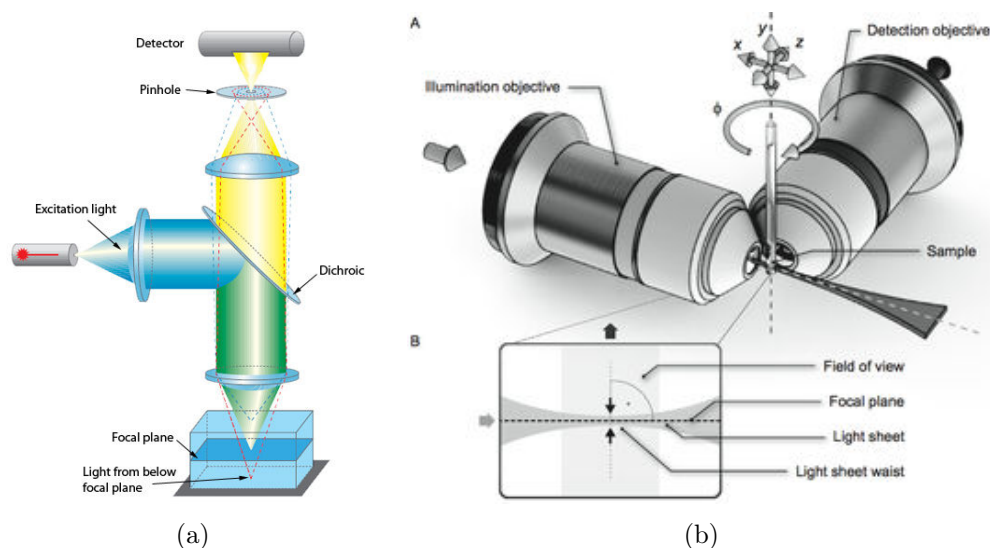


FIGURE 2.4 – Schémas de fonctionnement de microscopes confocaux. (a) Microscope confocal à balayage laser [Johnson 2006]. (b) Microscope à feuille de lumière (SPIM) [Weber 2014].

2.2.1.2 Microscopie multiphotonique.

Cette technique d'imagerie se base sur l'effet d'excitation coopérative multiphoton. L'absorption simultanée de deux photons par une molécule a été théoriquement prédite par Maria Göppert, prix Nobel de physique en 1931, mais c'est seulement en 1990, avec des développements technologiques tels que des sources laser à impulsions ultra-brèves, que le premier microscope biphotonique est mis en place [Denk 1990]. Si une description de son fonctionnement peut être trouvée dans [Pons 2005], notons deux principaux avantages que confère cette technique. D'abord, la longueur d'onde d'excitation se situe dans l'infra-rouge qui est une plage de plus grande transparence des tissus biologiques, ce qui permet d'imager des tissus plus en profondeur qu'avec les techniques d'imagerie précédemment introduites. Surtout, l'excitation multiphotonique se situe dans un volume focal très confiné (figure 2.5), avec comme conséquences une diminution conséquente du photoblanchiment et de la phototoxicité dus à l'acquisition. Les variantes les plus utilisées de la microscopie multiphotoniques sont les modes biphoton et triphoton. Si le microscope multiphoton est moins invasif et pénètre plus en profondeur que son concurrent confocal, son réglage est en revanche plus complexe et l'observation de tissus végétaux n'est pas encore bien maîtrisé avec cette méthode. Enfin, le prix d'un microscope multiphoton est très élevé par rapport à un confocal.

2.2.2 Marquage des structures

Lors d'études sur la morphogenèse d'organismes en développement, l'imagerie de tissus en développement se concentre principalement sur deux structures : les noy-

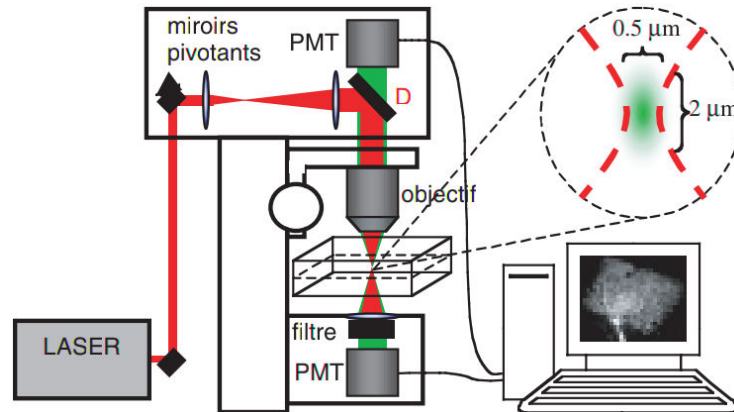


FIGURE 2.5 – Schéma de fonctionnement d'un microscope à fluorescence biphotonique [Pons 2005].

aux et les membranes des cellules. Lorsque ces structures ne sont pas naturellement fluorescentes (chlorophylle), elles doivent donc être marquées par des substances fluorescentes appelées fluorochromes. On distingue deux principales classes de marquage.

- Marquage vital : on immerge l'organisme dans une solution colorée par des molécules fluorescentes synthétiques qui se fixent directement sur des composants contenus exclusivement dans la structure visée. Les colorants Hoechst (Hoechst 33342 par exemple) ou DAPI sont utilisés pour le marquage de l'ADN contenu dans les noyaux de cellules. Les colorants styryliques (FM4-64 par exemple) sont quant à eux composés de molécules fluorescentes lipophiles qui s'attachent aux lipides présents dans les membranes cellulaires et permettent de les marquer. Les limitations du marquage vital sont le choix restreint des longueurs d'ondes pour l'imagerie, la pénétration parasite du colorant dans les membranes pour les colorants Hoechst et DAPI ou encore l'internalisation du colorant dans le cytoplasme des cellules pour les colorants styryliques.
- Marquage génétique : on introduit dans l'organisme un gène qui lui permettra de produire des protéines fluorescentes de manière endogène dans la structure désirée. On distingue deux types de marqueurs génétiques, les marqueurs biochimiques (principalement les isozymes) et les marqueurs moléculaires d'ADN. Contrairement au marquage vital, un organisme marqué génétiquement peut être observé à n'importe quel moment de son développement.

Il est également possible d'utiliser plusieurs marqueurs simultanément sur un organisme, ce qui permet d'acquérir des images à plusieurs canaux afin de superposer la visualisation de plusieurs structures [Day 2009].

2.3 Analyse d'images

L'émergence de techniques d'imagerie *in vivo* de plus en plus puissantes permet d'obtenir des bases de données de plus en plus grandes et nombreuses d'images de processus biologiques complexes. Par exemple, les microscopes à feuille de lumière peuvent acquérir chaque jour un grand nombre d'images représentant un volume dépassant les 10 téraoctets de données. Les quantités de volumes sont telles qu'il est désormais indispensable de développer des outils d'analyse d'images automatisés pour exploiter ces données biologiques d'une richesse sans précédent [Keller 2013]. L'analyse d'images biologiques est un thème de recherche très vaste et fait l'objet de nombreux travaux.

Selon la structure imagée (noyau ou membrane cellulaire), les outils de traitement et d'analyse d'image varieront mais suivent toujours le même schéma. Des processus de pré-traitement permettent d'améliorer le rapport signal-sur-bruit. Suivent des techniques de segmentation des cellules (noyaux seuls ou cellules entières) corrélées ou non avec des techniques de suivi des lignées cellulaires accompagnées d'annotations quantitatives. Ces étapes, qui constituent la partie de reconstruction spatio-temporelle de l'organisme en développement, sont indispensables pour pouvoir exploiter pleinement le potentiel des données acquises massivement avec les nouvelles techniques d'imagerie [Luengo-Oroz 2011].

2.3.1 Pré-traitements

Une image peut être sujette à un bruit induit par la procédure d'imagerie. Des pré-traitements dont le but est d'améliorer le rapport signal-sur-bruit existent par le biais de filtrages. Des logiciels tels que ND-Safir [Boulangier 2010] réalisent le pré-traitement d'images soumises à un bruit de type gaussien ou de Poisson. De manière générale, les classiques filtres gaussien et médian [Bloch 2005, Keller 2008, Long 2009] sont largement utilisés afin de lisser le bruit contenu dans le signal, mais des filtres plus sophistiqués tels que des équations aux dérivées partielles (EDP) [Kriva 2010], morphologie mathématique [Luengo-Oroz 2007, Fernandez 2010a], filtrages anisotropes [Mosaliganti 2010, Mosaliganti 2012] ou encore seuillage de transformées en ondelettes [Boutet de Monvel 2001] permettent de répondre à des besoins plus spécifiques.

L'optique des microscopes induit également un phénomène de convolution des images. La principale origine de ce phénomène est l'étalement de la réponse impulsionnelle optique, qu'on modélise mathématiquement par la fonction d'étalement du point (*Point Spread Function* ou PSF en anglais). Grâce à l'expérimentation, il est parfois possible de connaître l'opération de convolution. Dans ce cas, appliquer des opérateurs de déconvolution permet généralement d'obtenir une image plus nette et donc d'améliorer la résolution du système d'imagerie. Lorsqu'on procède à la déconvolution d'une image (on parle aussi de restauration), il est très important de modéliser le bruit qui entache le signal. En microscopie à fluorescence, le problème de déconvolution a fait l'objet de nombreuses études. Une vue d'ensemble des al-

algorithmes de restauration d'images en microscopie à fluorescence 3D est proposée dans [Sarder 2006].

Il existe aussi certains pré-traitements spécialement conçus pour une méthode d'acquisition en particulier. Par exemple, lors d'acquisitions d'images multiples du même objet sous différents angles (afin d'assurer une illumination plus homogène du tissu, voir [Krzic 2012]), il faut fusionner les différentes images obtenues après l'acquisition pour n'obtenir qu'une seule image.

2.3.2 Détection et segmentation des noyaux et des cellules

En biologie développementale, on distingue deux classes de méthodes de détection et de segmentation des images selon la structure observée (noyau ou membrane cellulaire).

2.3.2.1 Noyaux de cellules

Lorsqu'ils sont imagés, les noyaux de cellules apparaissent dans les volumes d'images comme des taches de haute intensité sur fond sombre (figure 2.6). De nombreuses techniques d'extraction et de segmentation des noyaux ont été développées et appliquées sur des données d'organismes modèles comme le *C. elegans* [Bao 2006, Dzyubachyk 2009], la drosophile [McMahon 2008, Supatto 2009, Amat 2014], le poisson-zèbre [Frolkovič 2007, Keller 2008] ou encore l'ascidie [Faure 2016].

De nombreuses méthodes de détection des noyaux font intervenir un seuillage d'image [Bao 2006, Al-Kofahi 2006, Keller 2008, Long 2009] auquel s'ajoutent des connaissances a priori de l'objet à segmenter (sa taille [Al-Kofahi 2006] ou sa forme [Bao 2006, Keller 2008] par exemple). D'autres méthodes utilisent un classifieur — de type forêt d'arbres décisionnels [Breiman 2001] par exemple — afin de classer chaque voxel selon son appartenance à un noyau [Sommer 2011]. [Keller 2010] fait quant à lui l'utilisation d'un filtre de Marr (laplacien de gaussienne) [Marr 1980]. [Frolkovič 2007] utilise une méthode sophistiquée basée sur les équations de convection-diffusion efficace sur les images fortement bruitées mais peu adaptée au traitement d'images contenant de nombreuses cellules à cause de son coût de calcul. L'identification des noyaux est rendue plus difficile lorsque les noyaux sont très proches les uns des autres. Une solution proposée par [Malpica 1997] est d'utiliser un algorithme permettant de discerner la frontière entre deux noyaux voisins, tel que l'algorithme de ligne de partage des eaux [Meyer 1991]. Cette liste de méthodes, à défaut d'être exhaustive, illustre la diversité des techniques de détection et de segmentation de noyaux. Enfin, des post-traitements permettent d'apporter des corrections aux segmentations obtenues.

2.3.2.2 Membranes cellulaires

Le marquage des membranes permet d'acquérir des volumes de données où les membranes sont intenses et le fond est sombre (figure 2.7). L'algorithme de

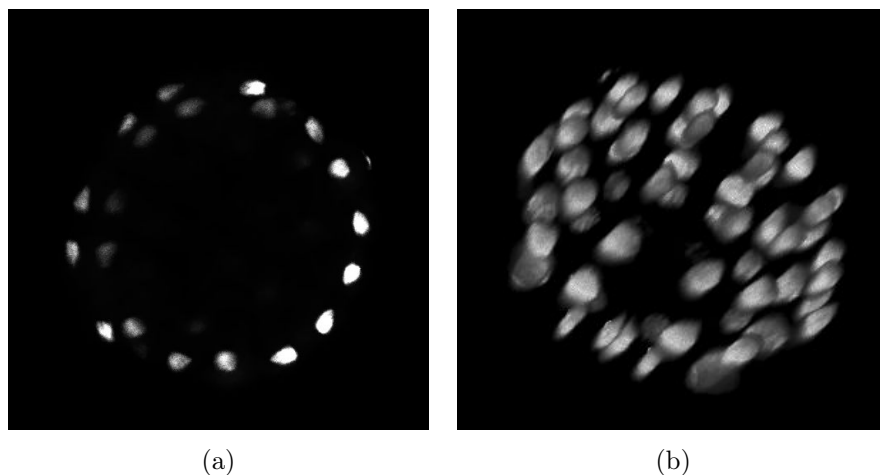


FIGURE 2.6 – Exemple d’image d’embryon d’ascidie obtenue par le marquage des noyaux. (a) Coupe 2D du volume de données. (b) Vue 3D de l’image.

ligne de partage des eaux [Meyer 1991] est particulièrement adapté à la segmentation de ce type de données et est assez largement utilisé dans la littérature en biologie développementale [Fernandez 2010a, Mkrtchyan 2011, Mosaliganti 2012, Khan 2014]. Cette méthode de segmentation considère une image en niveau de gris comme un relief topographique (figure 2.8) et calcule ladite ligne de partage des eaux de ce relief en simulant une inondation. Les régions d’intensité élevée de l’image forment des crêtes et les régions peu intenses forment des vallées appelées bassins versants (figure 2.9). Les voxels de l’image appartenant à un même bassin versant sont étiquetés à l’identique. On forme ainsi un partitionnement de l’image par des régions qui se rejoignent le long des crêtes. L’image d’origine peut contenir des structures parasites et du bruit, créant potentiellement des petites vallées qui engendreront de nombreuses sur-segmentations. Pour régler ce problème, un pré-traitement de l’image est généralement appliqué et on utilise une version de la méthode où l’on choisit les sources (graines) de l’inondation [Adams 1994]. Dans ce cas, il est nécessaire de proposer à l’algorithme des graines de bonne qualité. Pour la segmentation des cellules, cela revient à un problème de détection des cellules. Dans certains cas, lors d’une acquisition simultanée d’un canal « noyaux » et d’un canal « membranes cellulaires », on peut procéder à la détection des cellules à l’aide de l’image de leurs noyaux [Khan 2014]. Cette double acquisition est cependant plus difficile à obtenir d’un point de vue protocolaire. Selon les organismes développés, il est fréquent que seules les membranes soient imagées [Fernandez 2010a, Mkrtchyan 2011, Gelbart 2012, Mosaliganti 2012, Guignard 2014]. Dans ce cas, il est nécessaire d’implémenter une méthode de détection automatique de graines spécialement conçue pour le type de donnée que l’on traite.

La segmentation des membranes cellulaires peut aussi se faire par le biais d’autres méthodes. [Cilla 2015] propose une méthode de segmentation basée sur des modèles déformables via la détection des jonctions d’ancrage entre les cellules épithéliales

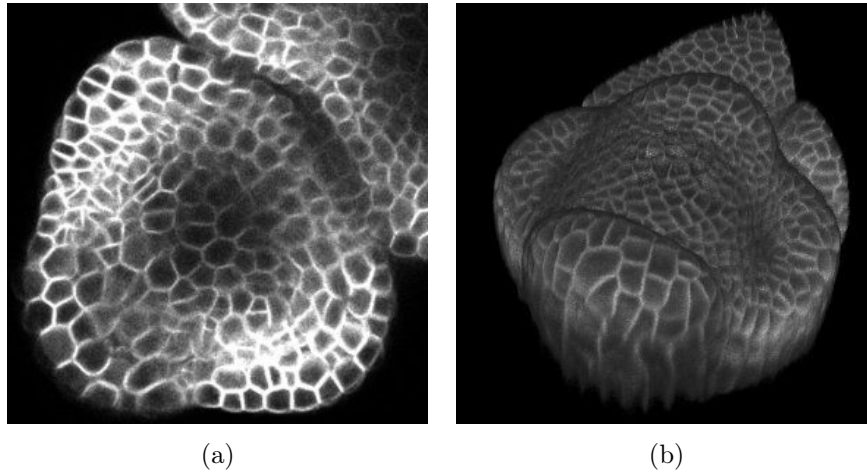


FIGURE 2.7 – Exemple d'image de méristème floral d'Arabidopsis obtenue par le marquage des membranes cellulaires. (a) Coupe 2D de l'image. (b) Vue 3D de l'image.

d'un tissu animal tel qu'une patte de drosophile. [Zanella 2010] propose l'utilisation de surfaces subjectives [Sarti 2002] et d'équations aux dérivées partielles, de même que [Mikula 2011], afin de segmenter les cellules d'embryons de poissons-zèbres. L'algorithme proposé par [Stegmaier 2016] repose sur une segmentation coupe par coupe (2D) de l'image puis sur une fusion des segmentations accompagnée par une technique d'optimisation combinatoire discrète pour corriger les segmentations finales en 3D. Une technique d'optimisation linéaire, testée sur des images de cione, permet d'effectuer la correction jointe de multiples sur-segmentations [Delibaltov 2013].

L'étape de segmentation des cellules est une tâche très complexe. Les méthodes de segmentation sont toujours développées afin de répondre à un besoin spécifique sur un type d'image, et une méthode plus adaptée qu'une autre pour un certain jeu de données ne le sera pas forcément pour d'autres données. Le choix d'une méthode de segmentation doit se faire en fonction de la situation : le type d'information biologique recherchée, le type d'acquisition, le type de tissu, la taille des données, des critères calculatoires spécifiques.

2.3.3 Lignées cellulaires

Les récentes techniques d'acquisition permettent d'imager des organismes en développement à une haute définition spatio-temporelle. La segmentation des cellules sur une image permet d'obtenir un modèle statique de l'organisme à un instant donné. Cependant, l'aspect temporel est primordial dans le but de numériser l'organisme en 4D ($3D+t$). Le suivi des lignées cellulaires (on parle aussi de filiation des cellules) au cours du développement d'un organisme permet de modéliser son aspect évolutif. Les algorithmes de suivi des cellules retracent l'arbre généalogique de chacune des cellules qui composent l'organisme en construisant une table de correspondances entre les cellules segmentées d'un ensemble d'images. Des règles de

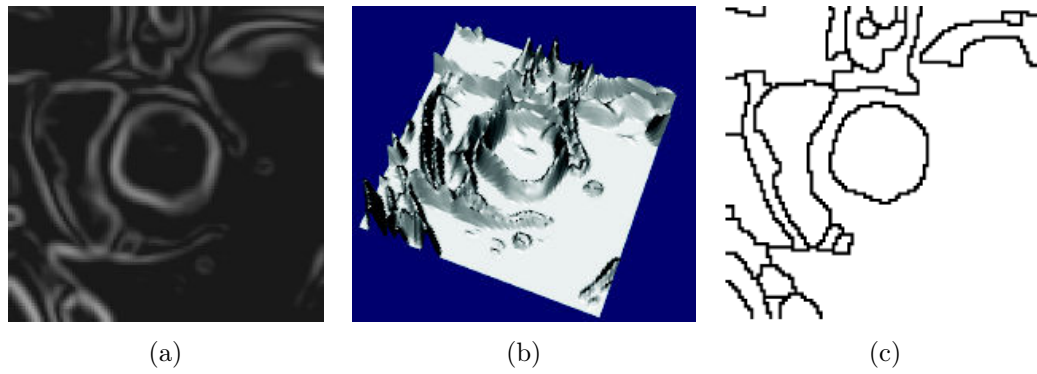


FIGURE 2.8 – Principe de base de l’algorithme de ligne de partage des eaux. (a) Image initiale en niveaux de gris. (b) Représentation du relief topographique de l’image. (c) Image segmentée. [Wikipedia 2015].

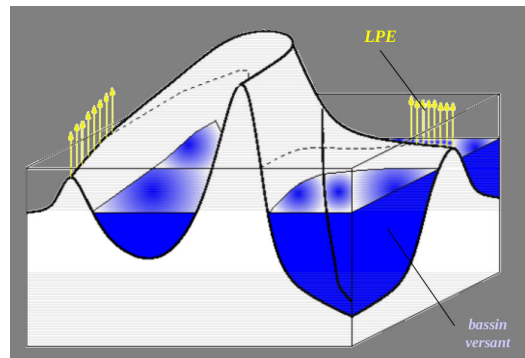


FIGURE 2.9 – Ligne de partage des eaux : l’ensemble des crêtes forme la LPE, l’ensemble des bassins versants forme les régions segmentées. [Manzanera 2007].

construction des correspondances cellulaires peuvent être exprimées lorsqu'elles ont un sens biologique. Par exemple, une règle classique est qu'une cellule a exactement un ancêtre à chaque instant, excepté lorsque des fusions de cellules peuvent survenir.

L'extraction de l'ensemble des lignées cellulaires d'un même organisme en développement offre la possibilité de réaliser des mesures quantitatives sur des déplacements d'ensembles de cellules (création d'organes), les évolutions de taille et de forme d'une même lignée, ou encore des mesures portant sur le cycle cellulaire (divisions cellulaires) [Fernandez 2010b].

Une première approche consiste à recalcr deux images successives d'une séquence temporelle, puis à rechercher l'antécédent de chaque cellule de la deuxième image dans la première image (algorithme du plus proche antécédent) [Bao 2006, Keller 2008]. Pour la mise en correspondance des cellules, une mesure de distance et/ou un critère de similarité entre cellules segmentées peuvent être utilisés. Pour recalcr deux images au sein d'une même séquence, de nombreuses méthodes de recalcr ont été développées [Sotiras 2013]. On citera en particulier l'algorithme de recalcr par blocs (block matching) [Ourselin 2000] qui consiste à itérativement estimer des appariements de blocs d'images, puis à calculer une transformation aux moindres carrés.

Le problème du calcul des lignées cellulaires peut aussi être considéré comme un problème d'optimisation [Fernandez 2010a, Cilla 2015]. Le problème s'exprime alors comme une fonction de coût à minimiser qui est égale à la somme des distances entre cellules appariées [Fernandez 2010a] ou comme un problème de flot de coût minimum (théorie des graphes) [Cilla 2015].

Les techniques de recherche du plus proche antécédent gèrent difficilement la possibilité d'avoir un champ de vue limité de l'organisme en développement (par exemple lorsqu'on image une plante), car des cellules peuvent alors entrer et sortir du champ de vue. Une autre difficulté apparaît lorsque la résolution temporelle de la série est trop faible pour assurer un fort recouvrement d'une même cellule entre deux images successives. Pour effectuer un suivi des lignées cellulaires de qualité, il faut donc privilégier une technique d'acquisition rapide comme la microscopie à feuille de lumière [Krzić 2012]. Enfin, les erreurs de segmentation représentent un facteur important d'erreur de suivi [Marcuzzo 2009].

Il est alors intéressant de considérer non-plus le problème de la segmentation des cellules comme un problème statique mais plutôt dynamique, c'est-à-dire de coupler la segmentation des cellules et leur suivi temporel. [Liu 2014] propose un algorithme qui considère de multiples segmentations de chacune des images d'une séquence et tente d'extraire la lignée cellulaire la plus probable pour chaque cellule de la première image de la série. Pour cela, il établit de manière exhaustive tous les appariements possibles selon le choix des segmentations et pondère les liens entre chaque paire de régions appariées selon le pourcentage de recouvrement entre ces régions. Cet algorithme est cependant très coûteux et ne peut être utilisé que sur des séries ayant peu d'images et de cellules. [Amat 2014] propose une méthode de reconstruction et de suivi des lignées cellulaires appliquée à des séries d'images d'embryons de drosophiles dont les noyaux sont marqués. Cette méthode se déroule

en trois étapes. D’abord, la taille des données est réduite en groupant les voxels des images en super-voxels. Puis l’utilisation de modèles déformables permet d’ajuster et de segmenter les noyaux des cellules. La segmentation obtenue à un instant donné est utilisée comme une initialisation pour le traitement de l’instant suivant. En plus d’aider le processus de segmentation, cette méthode fournit les correspondances entre les cellules segmentées des deux instants. Un traitement *a posteriori* permet d’extraire des informations morphologiques pour chaque cellule. Enfin, l’algorithme analyse pour chaque noyau détecté son voisinage pour identifier et éventuellement corriger les erreurs de reconstruction.

L’amélioration permanente des techniques et des protocoles d’acquisition des données offre des séquences spatio-temporelles d’une qualité toujours croissante. En y ajoutant les progrès à venir en termes de capacité d’utilisation des ressources informatiques (GPU, clusters, parallélisme), il est désormais envisageable de numériser entièrement le développement d’organismes à l’échelle cellulaire sur une longue durée. L’état de l’art propose un certain nombre de mesures permettant une description quantitative du développement d’organismes tels que l’arabidopsis [Fernandez 2010a], le drosophile [Blanchard 2009, Santella 2010, Amat 2014], le *C. elegans* [Santella 2010, Moore 2013], le poisson-zèbre [Blanchard 2009, Santella 2010] ou encore l’ascidie [Tassy 2006, Guignard 2015].

2.3.4 Quantification de l’expression des gènes

La capacité future de numériser et modéliser presque entièrement le développement d’un organisme en développement pose un nouveau défi. Il s’agit de corrélérer l’expression spatio-temporelle des gènes avec le mécanisme de contrôle qui spécifie les comportements et changements de forme des cellules au cours de la morphogenèse. Cela requiert, en plus de l’analyse quantifiée de séquences d’images, des méthodes offrant une solution pour différents problèmes :

- la fusion d’images multi-modales lorsque l’information est obtenue avec différentes techniques d’imagerie ;
- l’étude de séries temporelles plus longues en assemblant plusieurs petites séries ayant un recouvrement partiel [Guignard 2014] ;
- l’analyse statistique de populations et le recalage sur un atlas ;
- l’assemblage de données provenant de plusieurs individus dans un même espace (en raison du grand nombre de mécanismes de régulation d’expressions de gènes, tous ne peuvent pas être imagés simultanément [Luengo-Oroz 2011]).

L’un des principaux axes de recherche en analyse d’image en biologie développementale est donc le développement de stratégies de recalage permettant d’aligner à la fois la morphologie et la vitesse de développement de différents individus dont l’orientation spatiale et la morphologie peuvent varier. Le développement d’atlas 4D d’expressions de gènes, en plus d’outils d’extraction et d’analyse de lignées cellulaires, permettrait à terme une représentation multi-niveau des organismes en développement où la forme et la position de chaque cellule ainsi que l’expression des

gènes seraient décrits pour tous les instants du développement [Megason 2007].

Le recalage d'images [Zitová 2003] est un domaine de recherche actif trouvant de nombreuses applications en télédétection, vision par ordinateur et tout particulièrement en imagerie médicale, dont [Sotiras 2013] propose une étude bibliographique étayée.

En biologie développementale, les techniques de recalage sont généralement adaptées de l'imagerie médicale. [Castro 2009] propose une mesure de similarité basée sur l'information mutuelle pour le recalage spatial d'embryons de poissons-zèbres. [Luengo Hendriks 2006, Surkova 2013] et [Long 2009] proposent un recalage orienté objet appliqué respectivement à la drosophile et au *C. elegans*. [Fowlkes 2008] propose un algorithme de recalage spatial entre deux embryons de drosophiles au même stade de développement en détectant des équivalences entre noyaux de cellules. Il aligne ensuite ces noyaux à l'aide d'un processus itératif qu'il initialise à l'aide d'informations extraites comme l'axe antérieur-postérieur de l'embryon. Enfin, il propose un algorithme d'alignement (recalage) temporel d'une image de noyaux de cellules sur un pattern temporel. Dans un travail plus récent, [Guignard 2014] propose une méthode de recalage spatio-temporel d'embryons d'ascidies (*Phallusia mammillata*) marquées au niveau des membranes cellulaires et dont le schéma de développement est très stéréotypé. Cette méthode s'appuie sur la détection de régions d'intérêt spatio-temporelles qui sont les régions faisant apparaître les plus larges déformations. L'alignement entre deux séquences d'images d'embryons est alors réalisé en trouvant les appariements de régions d'intérêt optimaux au sens des moindres carrés de distance.

Malgré l'émergence de ces quelques méthodes de recalage d'organismes en développement, l'état de l'art actuel n'est pas suffisant pour permettre des analyses quantitatives de populations et de construire des atlas $3D+t$ d'expressions de gènes. Dans cette optique, le développement de nouvelles stratégies automatisées de recalage spatio-temporel d'individus revêt une grande importance et constitue l'une des perspectives majeures de l'analyse d'image en biologie développementale.

2.4 Plan de la thèse

Le travail présenté dans cette thèse recouvre essentiellement les aspects de segmentation d'images et de recalage inter-individus. L'objectif de mon travail s'inscrit dans le cadre de l'Inria Project Lab (IPL) Morphogenetics qui intègre des recherches transversales en biologie développementale et dont le but est de simuler la croissance des cellules et des tissus à différentes échelles. Dans ce sens, j'ai essentiellement travaillé sur des séquences d'images 3D marquées au niveau des membranes provenant de deux organismes modèles que sont l'ascidie *Phallusia mammillata* et le végétal *Arabidopsis thaliana*.

Dans le chapitre 3, j'introduis un outil de comparaison des segmentations. Cet outil nous sera utile à plusieurs fins. Il permettra de quantifier et de qualifier objectivement la qualité d'une segmentation en présence d'une vérité terrain. À défaut

de vérité terrain, il permettra une comparaison quantifiée entre deux segmentations d'un même objet, ce qui rend leur analyse plus aisée. Enfin, cet outil fournit un moyen de dresser un tableau de correspondance entre les régions segmentées de deux images.

Je propose ensuite dans le chapitre 4 une méthode générale de segmentation des cellules basée sur une technique originale de détection et de reconstruction des structures planaires. Cette technique s'appuie en particulier sur une technique de groupement perceptuel de l'information [Koffka 1935] permettant de regrouper des structures contenues dans une image selon des indices de proximité et de bonne continuation.

Le chapitre 5 propose une méthode automatisée de recalage inter-individus spécifique à l'embryon d'ascidie. Le développement de l'ascidie est très stéréotypé, c'est-à-dire qu'à un même stade de développement, il est possible d'identifier des correspondances de cellule à cellule entre deux embryons et que l'agencement global des cellules varie peu au sein d'une même espèce. La méthode que l'on propose s'appuie sur le détecteur de structures planaires proposé dans le chapitre 4.

Le chapitre 6 fournit une méthode de recalage spatio-temporel de séquences d'images de boutons floraux d'arabidopsis. Je propose une méthode de recalage spatial de boutons floraux qui s'appuie sur des hypothèses d'invariance de taille et de forme inter-individuelle. Cette partie présente enfin une méthode de recalage temporel entre séquences d'images de boutons floraux.

Finalement, le chapitre 7 fournit des éléments de discussion ouvrant sur des perspectives sur le travail réalisé au cours de ma thèse.

Comparaison de segmentations

Sommaire

3.1	Motivations	21
3.2	Cas de figure	24
3.3	Mesure de recouvrement de régions	25
3.3.1	Taux d'inclusion de l'érodé	25
3.3.2	Relation avec des indices symétriques	26
3.4	Matrices d'intersections de régions labellisées	27
3.5	Association des labels	28
3.6	Expériences	29
3.6.1	Données de test	29
3.6.2	Résultats	31
3.7	Conclusion	36

3.1 Motivations

La segmentation d'images est une opération de grande importance dans de nombreuses applications en vision par ordinateur, vidéo et image. En biologie du développement, l'étude de la morphogenèse des organismes en développement repose en partie sur l'analyse de la dynamique des cellules composant les tissus en phase de croissance. Pour toute analyse quantitative de cette dynamique, il est nécessaire de numériser les tissus imagés à l'échelle cellulaire. De nombreux outils sont mis en œuvre pour procéder à la segmentation des cellules dans le cadre de l'imagerie par fluorescence (voir section 4.1).

L'existence de multiples outils de segmentation d'images induit la question de l'évaluation de ces outils. Une revue des méthodes d'évaluation des segmentations est proposée dans [Zhang 2008]. Ces méthodes sont classées selon leur nature :

- Les évaluations subjectives sont les plus répandues et consistent à soumettre un résultat de segmentation au jugement d'un expert. Ces évaluations qualitatives sont fastidieuses et peuvent varier d'un expert à l'autre.
- Les évaluations analytiques évaluent les algorithmes de segmentation en se basant sur les types de stratégies employées, leur complexité et leurs besoins en ressources. Ces aspects sont indépendants des performances de segmentation finalement obtenues.

- Les évaluations supervisées comparent une segmentation test avec une segmentation de référence, généralement une vérité terrain réalisée manuellement. Elles permettent de quantifier la performance d’une segmentation avec précision mais il reste une part intrinsèque de subjectivité lors de la construction de la vérité terrain. De plus, cette tâche demeure parfois délicate et longue à réaliser. Il est possible de considérer l’évaluation de méthodes de segmentation sur des données synthétiques à la place de données réelles, l’avantage étant que la vérité terrain est alors parfaitement connue. En contrepartie, les données synthétiques peuvent être difficile à construire et ne garantissent pas une évaluation dans des conditions réalistes. Néanmoins, elles peuvent faciliter l’évaluation d’une méthode sur des critères ciblés.
- Les évaluations non-supervisées sont quantitatives et objectives. Elles ne nécessitent pas d’image référence, mais évaluent à la place une segmentation selon un ensemble de caractéristiques qu’elle est supposée vérifier. Ces évaluations sont intéressantes lorsqu’il est impossible ou trop complexe de produire une vérité terrain pour chaque image dont on souhaite évaluer la segmentation.

Le choix d’une méthode d’évaluation de segmentation dépend des critères que l’on souhaite mesurer. Deux mesures classiquement utilisées pour l’évaluation de la qualité de segmentations sont les indices de Jaccard [Jaccard 1901] — utilisé dans [Zijdenbos 1994, Dunham 2006] — et de Dice [Dice 1945] — utilisé dans [Fernandez 2010a] — (équations 3.1 et 3.2). Il s’agit d’indicateurs de similarité entre des segmentations d’un même objet, permettant d’évaluer la justesse des contours détectés d’une cellule a par rapport à une vérité terrain b . Ces mesures fournissent un score d’association symétrique entre les deux régions a et b ne faisant pas apparaître de paramètre de tolérance sur la position des frontières. Elles calculent le rapport entre le cardinal de l’intersection des ensembles comparés et le cardinal de l’union (Jaccard) ou la demi-somme des cardinaux (Dice) de ces ensembles :

$$\text{Jaccard}(a, b) = \frac{|a \cap b|}{|a \cup b|} ; \quad (3.1)$$

$$\text{Dice}(a, b) = 2 \frac{|a \cap b|}{|a| + |b|}. \quad (3.2)$$

Un nombre conséquent de mesures dérivées des indices de Jaccard et de Dice, reposant sur des théories probabilistes ou ensemblistes, sont apparues ces dernières années dans le but d’évaluer la qualité d’une segmentation test par rapport à une segmentation de référence (vérité terrain). On retrouve dans [Li 2009, Russell 2009] une stratégie d’évaluation de segmentation portée sur la mesure des ratios de vrais / faux positifs et de vrais / faux négatifs, tandis que [Korzynska 2007, Ko 2009, Srinivasa 2009] proposent une comparaison de la précision des formes des cellules extraites à l’aide de ratios basés sur les opérations d’union et d’intersection d’ensembles fortement inspirée des indices de Jaccard et de Dice. [Polak 2009] propose une métrique plus sophistiquée mesurant une erreur sensible aux sous-segmentations et aux sur-segmentations d’une segmentation A par rapport à une segmentation B ,

mais ne permettant pas de distinguer une sous-segmentation de A par rapport à B d'une sous-segmentation de B par rapport à A . [Held 2011] propose une comparaison de segmentations mêlant mesure de similarité et analyse statistique. Chacune de ces méthodes d'évaluation permet de quantifier la qualité d'une segmentation test par rapport à une segmentation de référence, au sens de la précision des formes extraites, à l'aide d'une mesure scalaire unique résumant la comparaison entre ces deux segmentations. À l'inverse, [Dima 2011] étudie l'utilisation d'un indice de similarité asymétrique dans le but de distinguer les cas de sur-extraction et de sous-extraction d'une segmentation test par rapport à une segmentation de référence.

Le travail présenté au chapitre 4 concerne la création d'un nouvel outil de segmentation générique de cellules de tissus épithéliaux. Pour évaluer les performances de cet outil, nous avons besoin d'une comparaison quantitative des segmentations obtenues par celui-ci avec celles obtenues par d'autres méthodes de segmentation [Fernandez 2010a, Mosaliganti 2012]. Dans notre cas, des vérités terrain sont obtenues par la correction manuelle d'une segmentation réalisée par l'une des méthodes de référence [Fernandez 2010a] dont on souhaite comparer les performances avec l'outil que nous avons développé. L'expertise apportée se situe donc au niveau de l'identification des cellules et l'évaluation d'une segmentation basée sur cette vérité terrain ne devrait pas porter sur la précision de la localisation des frontières des régions segmentées.

Dans ce chapitre, nous cherchons à définir une méthode de comparaison de segmentations de tissus épithéliaux permettant de répondre à la question « a-t-on les mêmes ensembles de cellules ? » entre deux segmentations A et B d'une même image. Dans notre problème, nous ne souhaitons donc pas évaluer la forme des cellules extraites mais la capacité d'identification des cellules d'une méthode de segmentation. Ce problème d'identification induit le besoin de mettre en correspondance des régions provenant de différentes segmentations, avec une nécessaire flexibilité sur la localisation des frontières des régions segmentées (figure 3.1). La mise en correspondance de régions segmentées est un problème asymétrique par nature. Par conséquent, les mesures de Dice, de Jaccard ainsi que leurs dérivées symétriques ne sont pas adaptées à cette problématique. Dans la suite de la section, nous proposons d'utiliser une paire d'indices asymétriques proche de [Dima 2011] permettant de mesurer la projection d'une région segmentée de A dans l'ensemble des régions de B et réciproquement. Ces indices ont pour but de mettre en correspondance des régions de A et B . L'interprétation de ces associations permet d'identifier les cas de détection similaire, de sous-détection, de sur-détection de cellules entre deux segmentations.

Notations. Une segmentation consiste en une image de labels où chaque label correspond à une structure (ici une cellule). Nous considérons dans la suite deux segmentations $A = \{a_i\}_{i \in [0, n]}$ et $B = \{b_j\}_{j \in [0, m]}$ d'une même image I . Les labels a_0 et b_0 correspondent aux labels de fond de leur segmentation respective. Les autres labels correspondent aux régions segmentées de I .

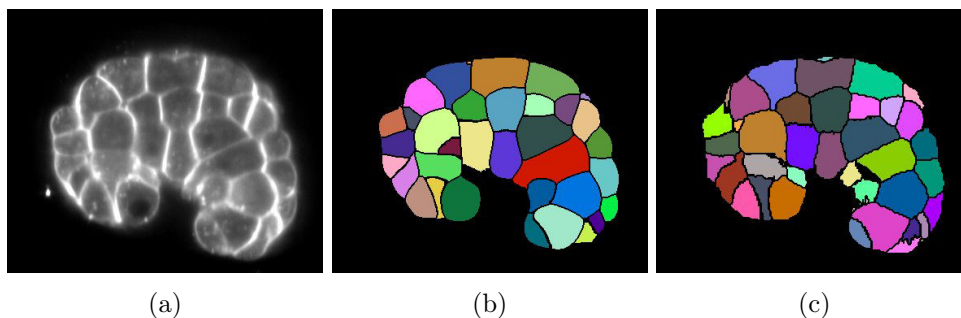


FIGURE 3.1 – Vues en coupe d’une image 3D d’embryon d’ascidie marquée au niveau des membranes cellulaires (a), de sa segmentation obtenue avec une méthode test (b) et de sa segmentation de référence (c). L’objectif de ce chapitre est de proposer un outil de comparaison de segmentation permettant d’établir les correspondances de cellules détectées entre les deux segmentations (b) et (c) et de quantifier les cas de segmentations communes (correspondance cellule à cellule) et de segmentations distinctes.

Organisation du chapitre. Nous proposons une méthode de comparaison de segmentations permettant d’identifier différents cas de figures que nous présentons en section 3.2. Une mesure de recouvrement de labels (section 3.3) nous permet de calculer des matrices d’intersection label à label (section 3.4). On extrait de la matrice d’intersection des associations de labels dans la section 3.5. Nous présentons nos expérimentations en section 3.6 puis concluons en section 3.7.

3.2 Cas de figure

Nous cherchons à déterminer les cas de figure suivants, illustrés en figure 3.2, lors de la comparaison de la segmentation A par rapport à la segmentation B :

- Correspondance bilatérale entre un label de A et un label de B . Dans la suite, nous nommons ce cas de figure une *bijection*.
- Correspondance entre un label de A et plusieurs labels de B . Ce cas de figure constitue à la fois une *sous-segmentation* de A par rapport à B et une *sur-segmentation* de B par rapport à A .
- Correspondance entre plusieurs labels de A et un label de B . Il s’agit d’une *sur-segmentation* de A par rapport à B et d’une *sous-segmentation* de B par rapport à A .
- Correspondance d’un label de A avec le *label de fond* de B .
- Correspondance entre plusieurs labels de A et plusieurs labels de B . Nous nommons ce cas de figure une *divergence* de segmentations.

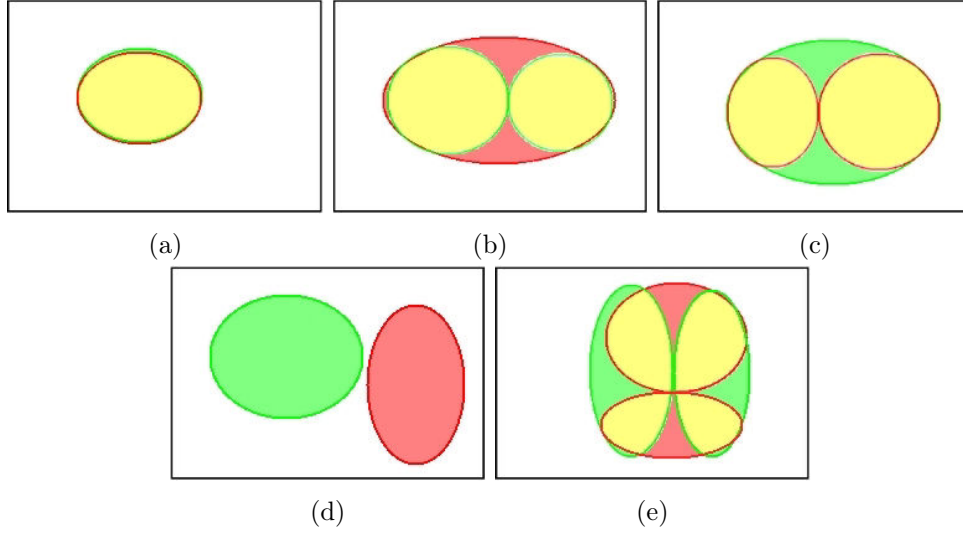


FIGURE 3.2 – Illustration des différents cas de figure attendus lors de la comparaison de deux segmentations A (rouge) et B (verte). (a) Bijection de labels. (b) Sous-segmentation de A par rapport à B . (c) Sur-segmentation de A par rapport à B . (d) Label de A dans le label de fond de B . (e) Divergence des segmentations de A et B .

3.3 Mesure de recouvrement de régions

3.3.1 Taux d'inclusion de l'érodé

Une région $a_i \in A$ peut être associée à une région $b_j \in B$ si ces régions s'intersectent. Une mesure d'intersection de régions peut être définie pour quantifier le taux d'inclusion d'une région dans l'autre et réciproquement. Le problème d'association de régions labellisées est asymétrique par nature. Il faut donc une mesure asymétrique pour permettre de distinguer une sur-segmentation d'une sous-segmentation. L'indice de similarité bivariant introduit par [Dima 2011] mesure le taux d'inclusion d'un label test a_i dans un label de référence b_j et réciproquement (équation 3.3). Cet indice est utilisé pour distinguer les cas de sur-extraction et de sous-extraction d'objets par rapport à une référence en terme de volume. Le taux d'inclusion de a_i par rapport à b_j s'interprète comme la probabilité conditionnelle d'un point de l'objet test a_i d'appartenir à b_j , et réciproquement :

$$\begin{aligned} ab_{ij} &= p(b_j|a_i) = \frac{|a_i \cap b_j|}{|a_i|}, \\ ba_{ji} &= p(a_i|b_j) = \frac{|a_i \cap b_j|}{|b_j|}. \end{aligned} \quad (3.3)$$

Afin d'assurer une tolérance sur la précision des frontières de régions que l'on compare, nous proposons d'introduire un paramètre d'érosion r dans l'équation 3.3 afin de considérer le taux d'inclusion ab_{ij}^r (respectivement ba_{ji}^r) d'un érodé a_i^r du label a_i dans le label b_j (respectivement d'un érodé b_j^r du label b_j dans le label a_i) :

$$\begin{aligned} \forall i, j \quad ab_{ij}^r &= p(b_j|a_i^r) = \frac{|a_i^r \cap b_j|}{|a_i^r|}, \\ ba_{ji}^r &= p(a_i|b_j^r) = \frac{|a_i \cap b_j^r|}{|b_j^r|}. \end{aligned} \quad (3.4)$$

Le paramètre r dénote le rayon d'érosion des régions, c'est-à-dire $a_i^r = a_i \ominus \mathcal{B}_r$ et $b_j^r = b_j \ominus \mathcal{B}_r$ où \mathcal{B}_r désigne une boule de rayon r . Ce paramètre permet de fixer la tolérance souhaitée sur la précision des frontières des régions à comparer.

Cette mesure, toujours comprise entre 0 et 1, vérifie les propriétés :

$$\begin{aligned} ab_{ij}^r &= 1 \Leftrightarrow a_i^r \subset b_j ; \\ ba_{ji}^r &= 1 \Leftrightarrow b_j^r \subset a_i. \end{aligned} \quad (3.5)$$

$$\begin{aligned} ab_{ij}^r &= 0 \Leftrightarrow a_i^r \cap b_j = \emptyset ; \\ ba_{ji}^r &= 0 \Leftrightarrow b_j^r \cap a_i = \emptyset. \end{aligned} \quad (3.6)$$

La mesure du taux d'inclusion de l'érodé est asymétrique. Elle permet d'établir des associations de labels avec une tolérance sur la localisation des frontières de régions grâce à son paramètre d'érosion.

3.3.2 Relation avec des indices symétriques

3.3.2.1 Indice de Dice

L'indice de Dice défini par l'équation 3.2 peut être redéfini à partir de la mesure du taux d'inclusion d'une région dans une autre définie par l'équation 3.4 dans le cas où le paramètre d'érosion r est fixé à 0 :

$$\begin{cases} \text{Dice}(a_i, b_j) &= 0 & \text{si } p(a_i|b_j) = 0 \text{ ou } p(b_j|a_i) = 0, \\ \frac{1}{\text{Dice}(a_i, b_j)} &= \frac{1}{2} \left(\frac{1}{p(a_i|b_j)} + \frac{1}{p(b_j|a_i)} \right) & \text{sinon.} \end{cases} \quad (3.7)$$

Cette relation permet d'établir que l'indice de Dice, extrêmement répandu dans les problèmes d'évaluation de qualité de segmentations, se déduit de la mesure introduite dans cette section. La considération de la paire de mesures d'inclusion $p(a_i|b_j)$ et $p(b_j|a_i)$ fournit donc naturellement une information plus riche que la mesure de Dice.

Note : Il est ainsi possible de définir par extension de l'équation 3.7 un indice symétrique de pseudo-Dice Dice_r , introduisant le paramètre d'érosion r , qui évalue la similarité entre deux segmentations d'une structure particulière avec une tolérance sur la localité de la frontière des segmentations :

$$\begin{cases} \text{Dice}_r(a_i, b_j) = 0 & \text{si } p(a_i|b_j^r) = 0 \text{ ou } p(b_j|a_i^r) = 0, \\ \frac{1}{\text{Dice}_r(a_i, b_j)} = \frac{1}{2} \left(\frac{1}{p(a_i|b_j^r)} + \frac{1}{p(b_j|a_i^r)} \right) & \text{sinon.} \end{cases} \quad (3.8)$$

Cet indice vérifie $\text{Dice}_r(a_i, b_j) = 1$ si et seulement si $a^r \subset b$ et $b^r \subset a$. Aussi, $\text{Dice}_r(a_i, b_j) = 0$ si et seulement si $a^r \cap b = \emptyset$ ou $a \cap b^r = \emptyset$.

3.3.2.2 Définition de distance

Dans [Dima 2011], une distance d entre deux régions segmentées a et b est proposée :

$$d(a, b) = \sqrt{(1 - p(a|b))^2 + (1 - p(b|a))^2}. \quad (3.9)$$

Si on réalise la projection du point de coordonnées $(p(a|b), p(b|a))$ dans \mathbb{R}^2 , cette distance s'interprète comme la distance euclidienne entre ce point et le point $(1, 1)$ correspondant à une parfaite correspondance entre les labels a et b .

Note : De même que pour l'indice de Dice, il est possible de définir par extension de l'équation 3.9 une pseudo-distance d_r paramétrée par une érosion de rayon r :

$$d_r(a, b) = \sqrt{(1 - p(a|b^r))^2 + (1 - p(b|a^r))^2}. \quad (3.10)$$

Avec cette mesure, $d_r(a, b) = 0$ signifie que $a^r \subset b$ et $b^r \subset a$. Cependant, l'égalité $a = b$ n'est pas forcément vérifiée. Il ne s'agit donc pas d'une distance mais d'un indice de similarité affichant une tolérance sur la position des frontières à l'instar du pseudo-Dice Dice_r .

Cependant, les indices symétriques Dice_r , d ou d_r conservent les désavantages que possèdent toutes les mesures symétriques de similarité de segmentations : elles ne permettent pas de distinguer les différents cas de segmentations non-communes.

3.4 Matrices d'intersections de régions labellisées

Nous considérons des segmentations faisant apparaître de multiples régions labellisées. Nous construisons à partir de la mesure donnée par l'équation 3.4 les matrices de taux d'inclusion de régions labellisées AB^r et BA^r :

$$AB^r = [ab_{ij}^r]_{i,j} ; BA^r = [ba_{ji}^r]_{j,i}. \quad (3.11)$$

On note que la somme des éléments de ces matrices le long d'une ligne donne 1 :

$$\begin{aligned} \forall i, \quad \sum_{j=0}^m ab_{ij}^r &= \sum_{j=0}^m p(b_j|a_i^r) = p(B|a_i^r) = 1 ; \\ \forall j, \quad \sum_{i=0}^m ba_{ji}^r &= \sum_{i=0}^m p(a_i|b_j^r) = p(A|b_j^r) = 1. \end{aligned} \quad (3.12)$$

Cas de figure	Caractérisation	Interprétation
$\{a_i\} \leftrightarrow \{b_j\}$	$ab_{ij}^r \simeq 1$ $ba_{ji}^r \simeq 1$	Bijection
$\{a_i\} \leftrightarrow \{b_{j_1} \cdots b_{j_q}\}$	$\sum_{l=1}^q ab_{ij_l}^r \simeq 1$ $ba_{j_l i}^r \simeq 1 \quad \forall l \in [1, q]$	Sur-segmentation de A / B
$\{a_{i_1} \cdots a_{i_p}\} \leftrightarrow \{b_j\}$	$ab_{i_k j}^r \simeq 1 \quad \forall k \in [1, p]$ $\sum_{k=1}^p ba_{j i_k}^r \simeq 1$	Sous-segmentation de A / B
$\{a_i\} \rightarrow \{b_0\}$	$ab_{i0} \simeq 1$	Label dans fond A/B
$\{a_0\} \leftarrow \{b_j\}$	$ba_{j0} \simeq 1$	Label dans fond B/A
$\{a_{i_1} \cdots a_{i_p}\} \leftrightarrow \{b_{j_1} \cdots b_{j_q}\}$	$\sum_{l=1}^q ab_{i_k j_l}^r \simeq 1 \quad \forall k \in [1, p]$ $\sum_{k=1}^p ba_{j_l i_k}^r \simeq 1 \quad \forall l \in [1, q]$	Divergence

TABLE 3.1 – Caractérisation et interprétation des cas de figure de comparaison de segmentations par l'étude du comportement des mesures d'intersection.

Le tableau 3.1 résume les comportements des mesures d'intersection attendus dans chaque cas de figure présenté en section 3.2. Le traitement particulier des labels de fond $a_0 \in A$ et $b_0 \in B$ permet de caractériser le cas de figure où un label de cellule se projette dans un label de fond.

3.5 Association des labels

On propose dans cette section d'associer chaque label d'une des segmentations à un unique label de l'autre segmentation en se basant sur les matrices d'intersections AB^r et BA^r précédemment calculées. L'associé d'un label $a_i \in A$ dans B se définit comme étant le label de B intersectant « le plus » le label a_i , c'est-à-dire le label maximisant le taux d'inclusion de a_i au sens de l'équation 3.4. L'association de $a_i \in A$ avec $b_j \in B$ se note $a_i \rightarrow b_j$:

$$a_i \rightarrow b_j \Leftrightarrow j = \arg \max_{j \in [0, m]} ab_{ij}^r. \quad (3.13)$$

Si on considère $b_j \in B$, on peut chercher l'ensemble des $a_i \in A$ qui sont « le plus » inclus dans b_j . On note cet ensemble A_{b_j} :

$$A_{b_j} = \{a \in A \mid a \rightarrow b_j\} \subset A. \quad (3.14)$$

De même, l'association de b_j à a_i se note $b_j \rightarrow a_i$, et on note B_{a_i} l'ensemble des labels de B qui s'associent à a_i :

$$b_j \rightarrow a_i \Leftrightarrow i = \arg \max_{i \in [0, n]} ba_{ji}^r. \quad (3.15)$$

$$B_{a_i} = \{b \in B \mid b \rightarrow a_i\} \subset B. \quad (3.16)$$

Par extension, nous notons les associations réciproques d'ensembles de labels $A' \subset A$ et $B' \subset B$:

$$A' \leftrightarrow B' \Leftrightarrow \left\{ \begin{array}{l} \forall a \in A', \exists b \in B' \mid a \in A_b \\ \forall b \in B', \exists a \in A' \mid b \in B_a \end{array} \right. . \quad (3.17)$$

Une association réciproque d'ensembles $A' \leftrightarrow B'$ est dite irréductible s'il n'existe pas de sous-ensembles $A'' \subset A'$ et $B'' \subset B'$ non vides tels que $A'' \leftrightarrow B''$.

Le cas particulier des associations aux labels de fond a_0 et b_0 est traité différemment des autres associations d'ensembles. Nous ne considérons pas d'association réciproque d'un label $a_i \in A$ avec b_0 (respectivement d'un label $b_j \in B$ avec a_0) mais seulement une association unilatérale $a_i \rightarrow b_0$ (respectivement $b_j \rightarrow a_0$). Autrement dit, nous ne calculons pas les associations de b_0 vers A ni de a_0 vers B .

Le calcul des associations de labels données par les équations 3.13 et 3.15 permet de déduire l'ensemble des associations réciproques irréductibles d'ensembles. Ces associations, caractérisées par le cardinal des ensembles associés, peuvent ainsi être interprétés d'après le tableau 3.1. On peut alors quantifier les segmentations qui se correspondent (bijections), les sur-segmentations et sous-segmentations, les détections de labels dans le fond et les cas de divergence de segmentations de A par rapport à B et de B par rapport à A . Des illustrations sont données en figure 3.3.

De plus, des statistiques de recouvrement volumiques entre les labels associés peuvent aussi être calculées, comme par exemple la proportion du volume total de la segmentation A en bijection avec la segmentation B et inversement, ou encore le taux de recouvrement moyen et son écart-type entre les labels en bijection d'une segmentation par rapport à l'autre.

3.6 Expériences

3.6.1 Données de test

Nous avons réalisé la comparaison de deux segmentations d'une image de méristème floral obtenue par microscopie confocale (figure 3.4a), ainsi que deux segmentations d'une image d'embryon d'ascidie acquise par microscopie à feuille de lumière (figure 3.5a). Les deux images sources sont marquées au niveau des membranes cellulaires. Elles font apparaître un réseau dense de cellules 3D séparées par leurs membranes et dont les segmentations forment des partitionnements denses des images associées.

Les segmentations de l'image de méristème floral sont obtenues par l'application de la méthode MARS [Fernandez 2010a] (figure 3.4b) et de la méthode de segmentation proposée dans le chapitre 4 (figure 3.4c). Ni l'une, ni l'autre de ces segmentations ne constitue une vérité terrain. Nous proposons de comparer ces deux segmentations dans le but de montrer les capacités de notre méthode de comparaison, à la fois d'identification des associations de cellules, de quantification des différences de segmentation et de visualisation de ces dernières.

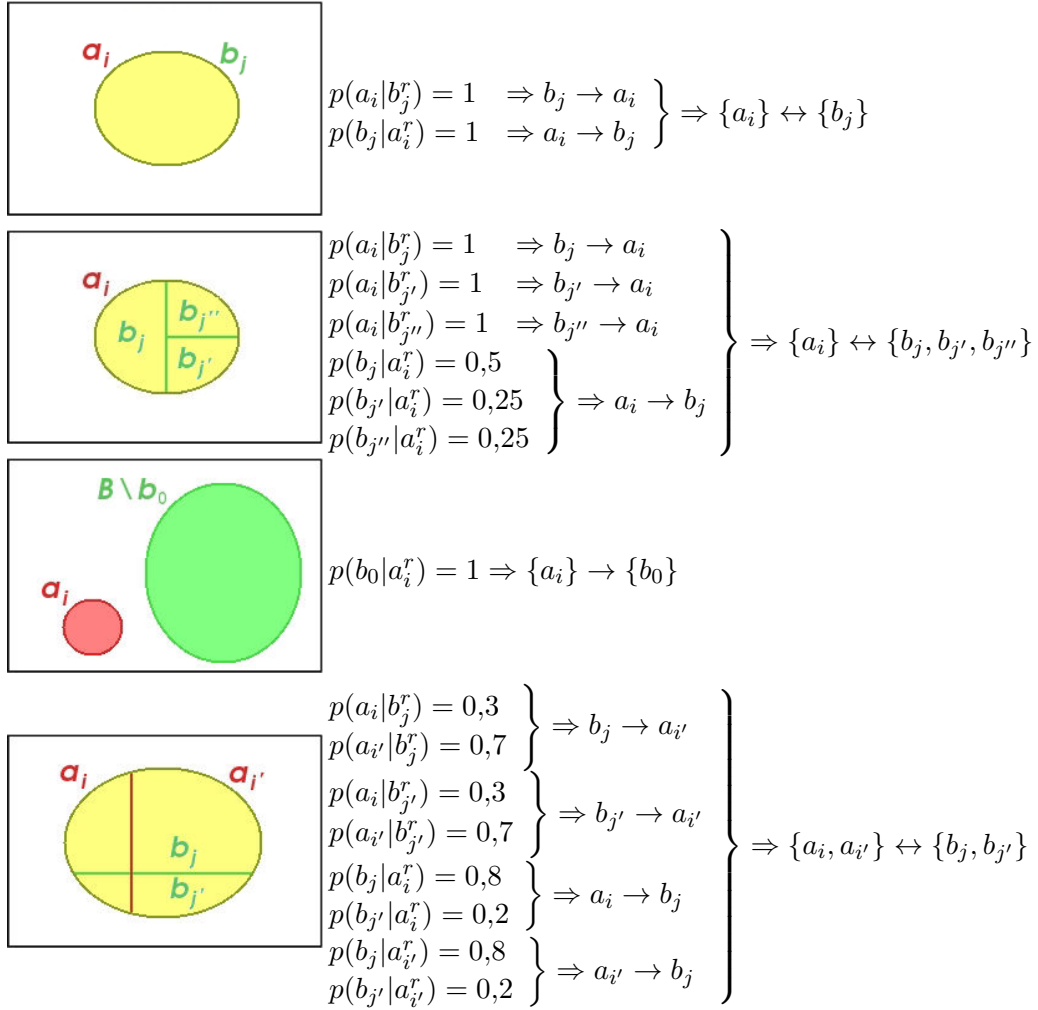


FIGURE 3.3 – Exemples de mesures de taux d'inclusion et d'association de régions labellisées dans différents cas de figure. Ligne 1 : bijection. Ligne 2 : sous-segmentation de A par rapport à B. Ligne 3 : label de A dans le fond de B. Ligne 4 : divergence entre segmentations.

Mesure	<i>A</i>	<i>B</i>
Nombre de labels	1354	1759
Bijections	933	933
Sous-segmentations	48	130
Sur-segmentations	280	175
Dans le fond	1	98
Divergences	90	421
Taux moyen de recouvrement	0,907656	0,891585
Écart-type	0,054888	0,087522
Taux de volume en bijection	0,479486	0,469699

TABLE 3.2 – Comparaison quantifiée des segmentations *A* et *B* présentées en figure 3.4 pour $r = 2$.

Pour la seconde expérience, nous comparons une segmentation test (figure 3.5b) avec une segmentation de référence où l’identification des cellules a été expertisée (figure 3.5c). Dans cette étude, nous testons l’influence du paramètre d’érosion r sur les appariements et les taux de recouvrement calculés entre les cellules (figure 3.6).

3.6.2 Résultats

La comparaison des segmentations de l’image de méristème a été paramétrée par une érosion $r = 2$. Les segmentations, nommées *A* et *B*, présentent respectivement 1354 et 1759 cellules. Le tableau 3.2 décrit une liste de mesures quantifiant les segmentations communes et non-communes entre *A* et *B*. Si une analyse au cas par cas est nécessaire pour interpréter chaque différence de segmentation entre *A* et *B*, on note une tendance de la segmentation *A* à mal détecter les cellules présentes dans les couches profondes du méristème (zone faiblement contrastée) qui se traduit par des extractions de régions difformes et de grande taille, tandis que *B* semble proposer une segmentation de meilleure qualité dans cette zone de l’image. Cette tendance se retrouve dans le tableau avec la mesure du nombre de régions détectées en divergence : 90 régions de *A* correspondent à 421 régions de *B*, ce qui montre que *B* a tendance à proposer un partitionnement plus fin de l’image d’origine que *A* dans ces zones. On note également la possibilité de convertir l’information liée à la cellule en information volumique sur les images segmentées (statistiques portant sur les taux de recouvrement, sur les quantités de volumes en bijection, sur-segmentation, etc.). L’identification des associations de régions permet une visualisation du résultat de comparaison entre les deux segmentations (figure 3.4). Cette visualisation met en évidence les zones de l’image dont les résultats de détection des cellules diffèrent entre deux méthodes, ce qui facilite considérablement une éventuelle correction manuelle de segmentation. Un contrôle visuel 3D des mises en correspondance de cellules montre la fiabilité de la méthode de comparaison de segmentation que nous proposons sur cette image.

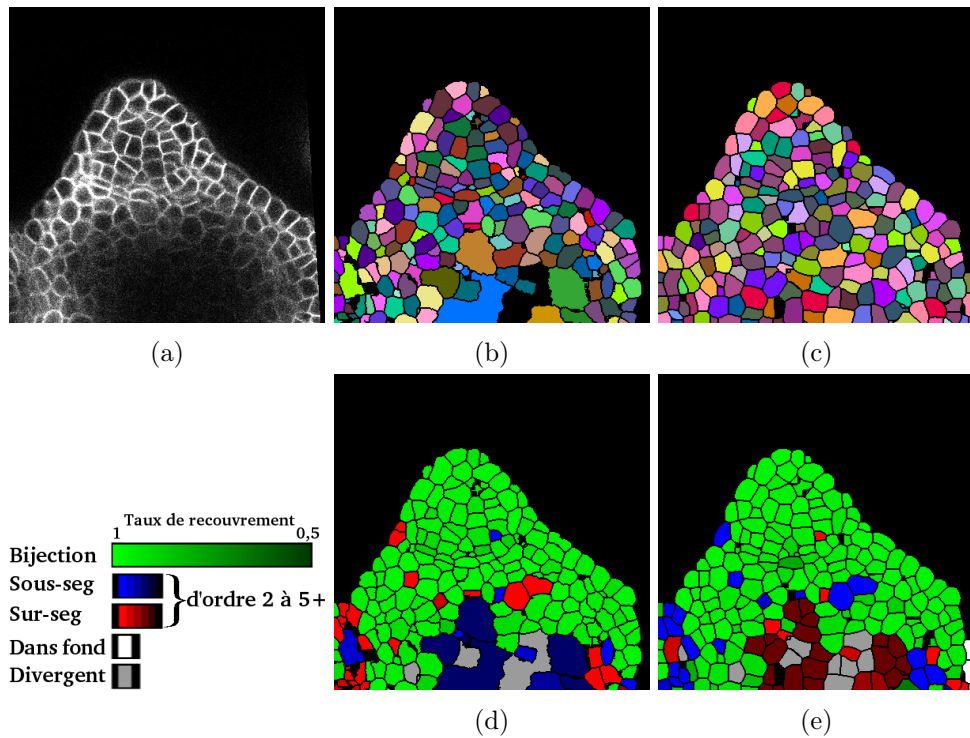


FIGURE 3.4 – Vues en coupe d’une image d’intensité I de méristème floral (a) et de deux segmentations A (b) et B (c) obtenues à l’aide de différentes méthodes (respectivement la méthode MARS [Fernandez 2010a] et la méthode décrite dans le chapitre 4). La coloration des régions de A (d) et B (e) correspond aux cas de figure détectés par la méthode de comparaison proposée ($r = 2$). Dans la légende, les ordres de sous-segmentation et de sur-segmentation réfèrent au nombre de régions d’une segmentation associés à une région de l’autre segmentation.

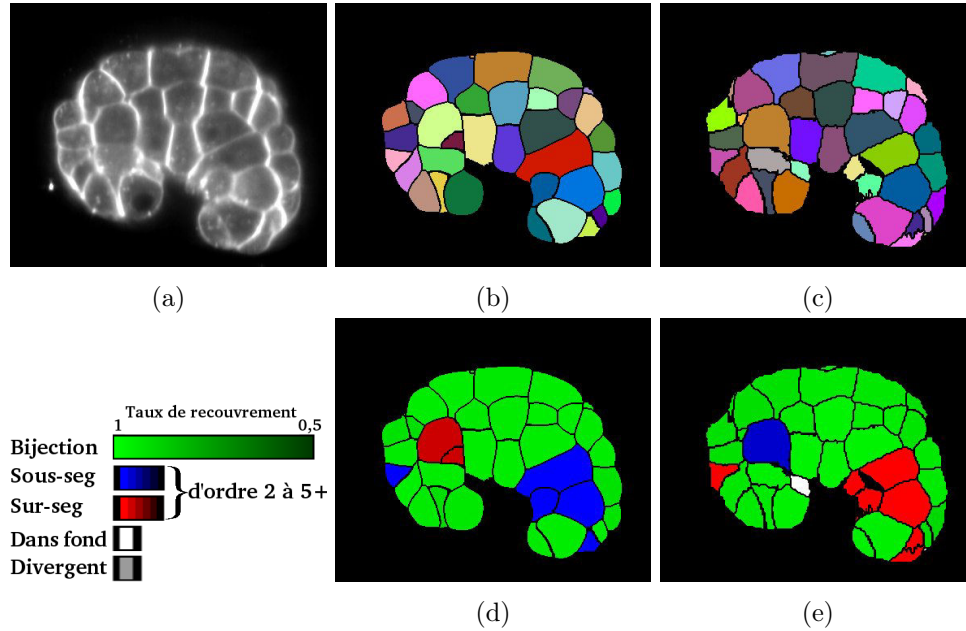


FIGURE 3.5 – Application de la mise en correspondance de régions provenant de différentes segmentations sur une image d’embryon d’ascidie. (a) Image originale (vue en coupe). (b) Segmentation A (méthode décrite au chapitre 4). (c) Segmentation B (méthode MARS [Fernandez 2010a]). (d,e) Coloration des régions segmentées de A et B selon le résultat des mises en correspondance de régions pour $r = 2$.

Dans notre seconde étude, les segmentations A et B sont respectivement composées de 166 et 162 régions labellisées. Le résultat de l’association des labels de A et de B est le même pour un paramètre d’érosion allant de 0 à 7, ce qui traduit une consistance appréciable dans les résultats de mise en correspondance de cellules. Les associations de labels commencent à varier pour $r \geq 8$. Il est naturel d’observer ces variations pour de grandes valeurs d’érosion puisqu’on ne garantit alors pas la préservation des régions les plus petites ou étroites. Pour une érosion $r \leq 7$, la comparaison indique que 133 cellules de A et de B sont en bijection (figure 3.5). Le choix du paramètre d’érosion a une incidence sur les taux de recouvrement mesurés entre les labels érodés a_i^r (respectivement b_j^r) et les segmentations pleines b_j (respectivement a_i) (figure 3.6). Sur la figure 3.7, nous montrons la variation de la distribution des taux de recouvrement ab_{ij}^r et ba_{ji}^r en fonction de $r \in [0, 7]$ pour les labels de A et B en bijection. On note qu’en l’absence d’érosion ($r = 0$), la distribution des taux de recouvrement de labels de A vers B (d’une moyenne supérieure à 90%) est globalement supérieure à celle des labels de B vers A (environ 86% de moyenne). Cette différence s’explique par le fait que les régions de A sont généralement moins volumineuses que celles de B . Pour une érosion $r = 4$, les taux de recouvrement moyens de A vers B et de B vers A sont supérieurs à 99%. Observer l’évolution des distributions de taux de recouvrement mesurés selon r renseigne sur l’incertitude portant sur la localisation des frontières de cellules selon la méthodologie employée.

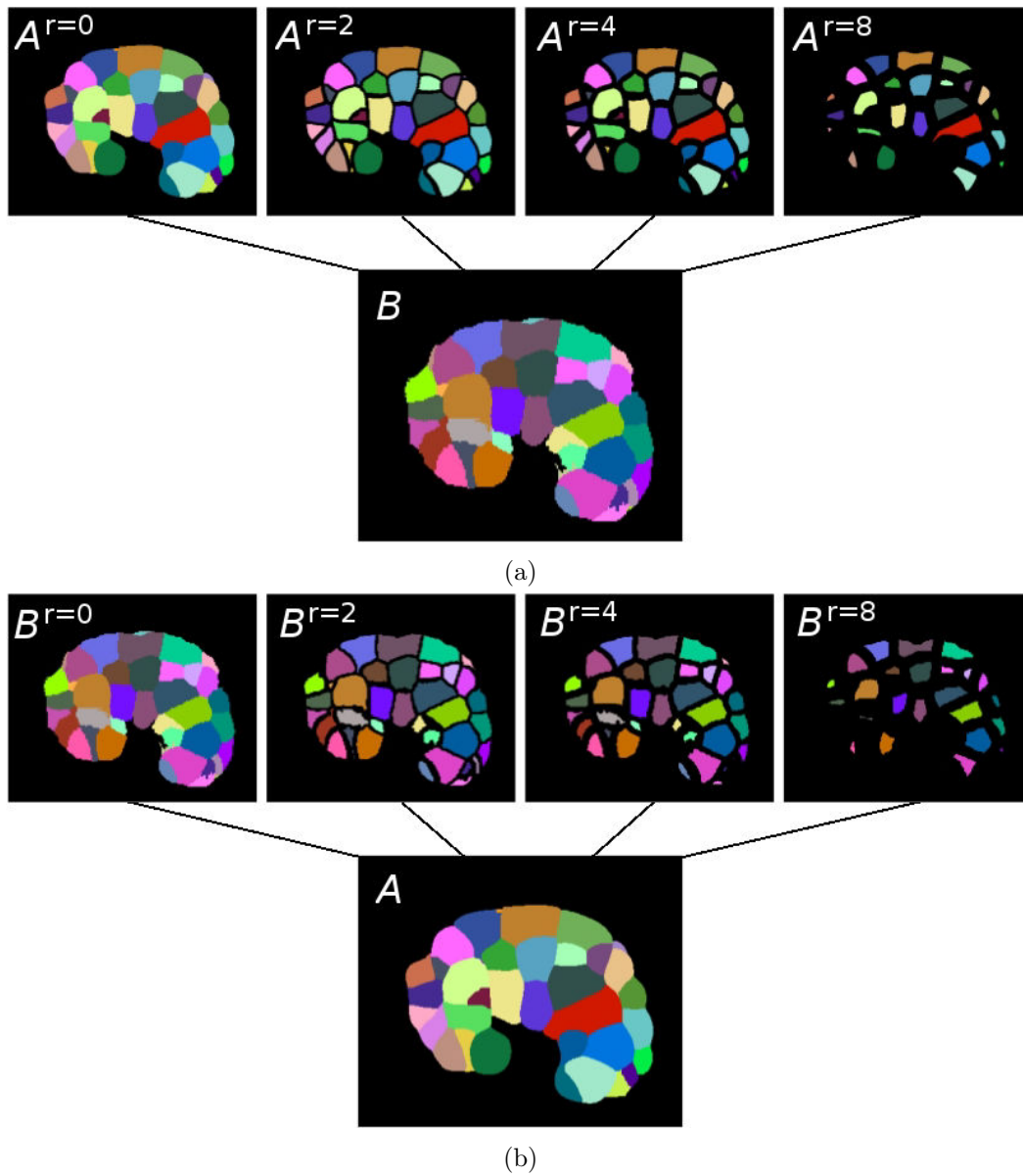


FIGURE 3.6 – Vue en coupe des segmentations comparées A et B pour un paramètre d'érosion $r = 0, 2, 4, 8$. (a) Comparaison de l'érodé A^r vers B . (b) Comparaison de l'érodé B^r vers A .

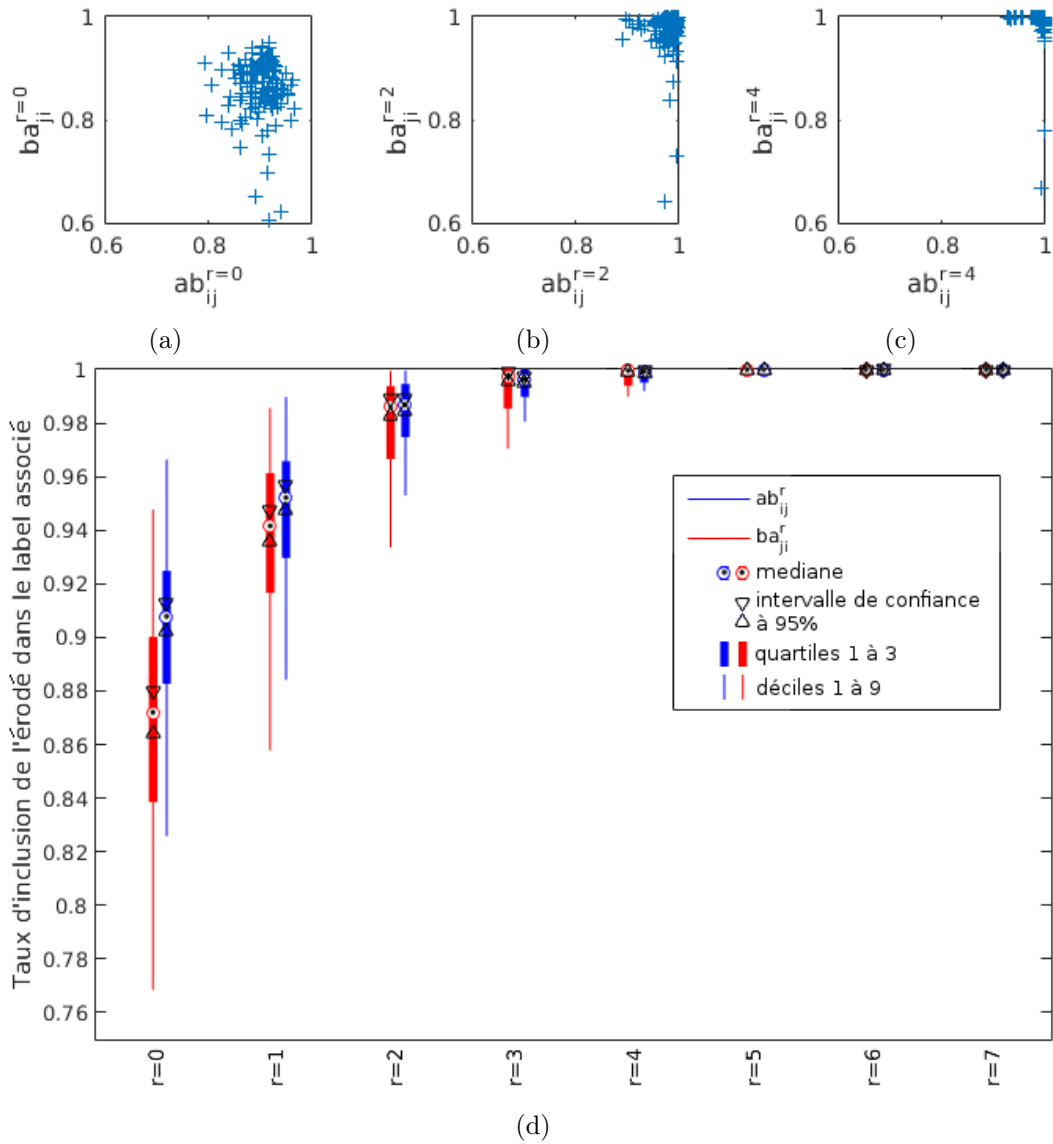


FIGURE 3.7 – Étude de l'influence du paramètre d'érosion r sur les taux de recouvrement calculés ab_{ij}^r et ba_{ji}^r entre labels en bijection, c'est-à-dire tels que $\{a_i\} \leftrightarrow \{b_j\}$, pour $r \leq 7$. (a-c) Distribution des taux de recouvrements de labels en bijection pour $r = 0, 2, 4$. Chaque croix est associée à une paire de labels en bijection (a_i, b_j) et est affichée aux coordonnées (ab_{ij}^r, ba_{ji}^r) . (d) Dispersion des taux de recouvrement des labels en bijection en fonction de l'érosion r (représentation en boîte à moustache).

3.7 Conclusion

Nous avons présenté dans ce chapitre une méthode de comparaison de segmentations que nous avons expérimentée et validée sur des images de tissus épithéliaux provenant d'organismes en développement (méristèmes floraux et embryons d'ascidies).

La méthode proposée permet de quantifier les similitudes et les différences dans les résultats de segmentation d'images grâce à l'identification des segmentations communes, des sous-segmentations et sur-segmentations de l'une des segmentations par rapport à l'autre et des régions faisant apparaître des divergences de segmentation.

Les applications de la méthode proposée sont d'abord l'évaluation supervisée d'une technique de segmentation grâce à une vérité terrain. En l'absence de vérité terrain, la comparaison de segmentations provenant de différentes méthodes permet de mettre en avant les régions faisant apparaître des différences de segmentations, à des fins de facilitation de la correction manuelle de segmentations. Enfin, la méthode de comparaison proposée permet également d'établir le lignage cellulaire entre deux segmentations d'images successives d'un même organisme, à condition de recalibrer les images dont on souhaite comparer les segmentations au préalable.

Dans la suite de ce manuscrit, nous utilisons la méthode de comparaison de segmentations proposée dans ce chapitre pour évaluer les segmentations obtenues à l'aide de la méthodologie présentée au chapitre 4. Dans le chapitre 6, nous nous appuyons aussi sur cette méthode pour établir des correspondances cellule à cellule entre deux images segmentées successives d'un même méristème floral en développement.

Segmentation de membranes cellulaires

Sommaire

4.1	Introduction	37
4.2	Fonction de planéité	39
4.2.1	Introduction	39
4.2.2	Modèle planaire	43
4.2.3	Fonction de réponse	43
4.3	Binarisation des membranes	45
4.3.1	Extraction des plans centraux	45
4.3.2	Inhomogénéité de l'intensité des membranes	47
4.3.3	Fonction de seuillage adaptatif	47
4.3.4	Histogrammes axiaux	48
4.3.5	Seuils axiaux	49
4.4	Reconstruction d'une image d'intensité par vote de tenseurs	50
4.4.1	Introduction	50
4.4.2	Les votes de tenseurs	51
4.5	Segmentation des cellules	57
4.5.1	Détection de graines	58
4.5.2	Ligne de partage des eaux	58
4.6	Expériences et résultats	58
4.6.1	Données	58
4.6.2	Évaluation de la binarisation des membranes	59
4.6.3	Évaluation du pré-traitement	61
4.6.4	Saillance membranaire	64
4.7	Conclusions et perspectives	65

4.1 Introduction

L'analyse de l'embryogenèse des plantes et des animaux constitue un axe de recherche majeur de la biologie développementale. Un des principaux défis à relever est de décrire l'évolution de la forme des organismes comme une conséquence de leur activité génique [Munro 2006, Traas 2009]. L'imagerie possède un rôle essentiel dans

l'observation spatio-temporelle d'organismes en développement pour l'étude et la quantification du comportement individuel et collectif des cellules qui les composent.

L'émergence récente de techniques d'imagerie 3D *in vivo* non-invasives rend aujourd'hui possible d'enregistrer la croissance d'animaux et de plantes avec une excellente résolution spatiale et une résolution temporelle suffisante pour permettre le suivi des déformations cellulaires au cours du temps [Keller 2013]. Les quantités de données générées par un microscope à fluorescence à illumination par feuille de lumière [Krzić 2012] peuvent atteindre plusieurs téraoctets par jour d'acquisition, et l'étude de différences phénotypiques nécessite d'imager des populations. Dans ce contexte, il est indispensable de posséder des outils d'analyse d'image par ordinateur aussi automatisés que possible.

L'analyse doit dépendre de ce qu'on cherche à observer. Selon les applications, les noyaux de cellules, les membranes ou murs cellulaires ou le cytoplasme peuvent être étudiés séparément ou conjointement. Les techniques de segmentation sont nombreuses et généralement très spécifiques à un type de données particulier. Un simple seuillage [Otsu 1979], des méthodes de contour actif [Cheng 2009], des méthodes de segmentation par coupe de graphe (graph-cut) [Qi 2014] et l'algorithme de ligne de partage des eaux [Malpica 1997] sont des méthodes communément utilisées pour procéder à la segmentation de noyaux de cellules. Pour l'analyse des interactions entre les cellules qui composent des tissus épithéliaux, l'analyse de la forme des cellules nécessite la segmentation des membranes cellulaires. Pour cela, des méthodes spécifiques ont été développées. La nature des images de cellules marquées au niveau des membranes s'adapte très bien à une approche de segmentation par ligne de partage des eaux marquée [Adams 1994], dont la bibliographie fait état de nombreux recours [Fernandez 2010a, Mkrtchyan 2011, Mosaliganti 2012, Khan 2014]. On retrouve également d'autres approches de segmentation basées sur des équations aux dérivées partielles [Zanella 2010], des surfaces subjectives [Sarti 2002], des approches de segmentation coupe par coupe de l'image [Stegmaier 2016] ou encore un problème d'optimisation linéaire qui effectue une correction jointe de plusieurs sur-segmentations [Delibaltov 2013].

Le travail proposé dans ce chapitre est motivé par le besoin d'outils dédiés à la segmentation d'images de tissus épithéliaux obtenues par microscopie à fluorescence. Pour cette application, les organismes imagés sont marqués au niveau de leurs membranes cellulaires. Lorsque les données traitées sont de grande taille (résolution, nombre de cellules par image, nombre d'images traitées), le besoin d'une solution au faible coût algorithmique restreint le choix du type d'approche. Dans ce contexte, la ligne de partage des eaux est l'algorithme de choix pour réaliser la segmentation de ce type d'image. Cependant, le bruit et les artefacts dus au procédé d'imagerie ainsi que les défauts de fluorescence sont sources d'erreurs de segmentation. Pour les éviter, l'étape de segmentation des cellules est toujours précédée d'un pré-traitement, généralement un filtrage plus ou moins sophistiqué, qui permet d'améliorer la qualité du signal en vue de sa segmentation. Récemment, des stratégies de filtrage spécifiques ont été développées pour augmenter le contraste des structures planes contenues dans une image [Mosaliganti 2010] et construire

une nouvelle image en niveaux de gris dans laquelle les structures planes sont propagées [Mosaliganti 2012]. L'image d'intensité reconstruite est alors segmentée par l'algorithme de la ligne de partage des eaux. Cette chaîne de traitement, nommée ACME (*Automated Cell Morphology Extractor*) par son auteur, a démontré sa capacité à segmenter des cellules dans un contexte difficile où les techniques de segmentation classiques sont mises à défaut. Un défaut majeur d'ACME est le coût computationnel de son étape de propagation des structures planes. Cette étape est réalisée par une approche de groupement perceptuel [Koffka 1935], nommément le vote de tenseurs [Medioni 2000].

Dans ce chapitre, nous proposons une chaîne de traitement automatisée conceptuellement voisine d'ACME [Mosaliganti 2012] pour la détection et la segmentation des cellules de tissus épithéliaux. Cette chaîne de traitement se décompose en deux principales étapes, la reconstruction d'une image d'intensité au signal rehaussé au niveau des membranes et la segmentation des cellules de l'organisme en partant de l'image reconstruite. Les principales contributions de nos travaux portent sur l'élaboration d'un nouveau filtre planaire basé modèle et d'une stratégie permettant de réduire le coût de calcul de l'étape de vote de tenseurs.

Dans la section 4.2, nous présentons un filtre planaire basé sur des informations dérivatives du premier et du second ordre. Ce travail est adapté d'un filtre linéique proposé par [Krissian 2000]. La section 4.3 présente une stratégie de binarisation des structures planaires réalisée par seuillage adaptatif selon l'orientation des structures. Cette stratégie est motivée par la nature anisotrope de certaines techniques d'imagerie par fluorescence [Krzic 2012] à l'origine de fortes variabilités de contraste des structures imagées selon leur orientation spatiale (figure 4.1). La section 4.4 présente une stratégie de reconstruction d'images par vote de tenseurs adaptée des travaux de [Mosaliganti 2012]. En section 4.5, nous présentons la méthode de segmentation des cellules basée sur une ligne de partage des eaux marquée. La stratégie de validation de la méthode de segmentation et les tests réalisés sur des données réelles (embryons de *Phallusia mammillata* et méristèmes d'arabidopsis) sont présentés en section 4.6. Nous proposons enfin une conclusion en section 4.7.

4.2 Fonction de planéité

4.2.1 Introduction

Différentes techniques de filtrage ont été développées en vue d'éliminer le bruit et d'augmenter le contraste des structures observées en imagerie biomédicale. Ces filtres peuvent être basés sur les propriétés photogrammétriques et/ou structurelles des objets à filtrer. Pour le filtrage de membranes, on trouve des méthodes de type détection de symétrie de phase [Kovesi 1997], des techniques de morphologie mathématique [Luengo-Oroz 2007], des filtres par ondelettes [Boutet de Monvel 2001] ou encore des filtres linéaires [Marcuzzo 2009]. Des filtres de diffusion anisotrope basés sur le modèle de [Kriva 2010] ont été plus récemment développés dans le but de débruiter les images tout en rehaussant le signal des membranes. Ces fil-

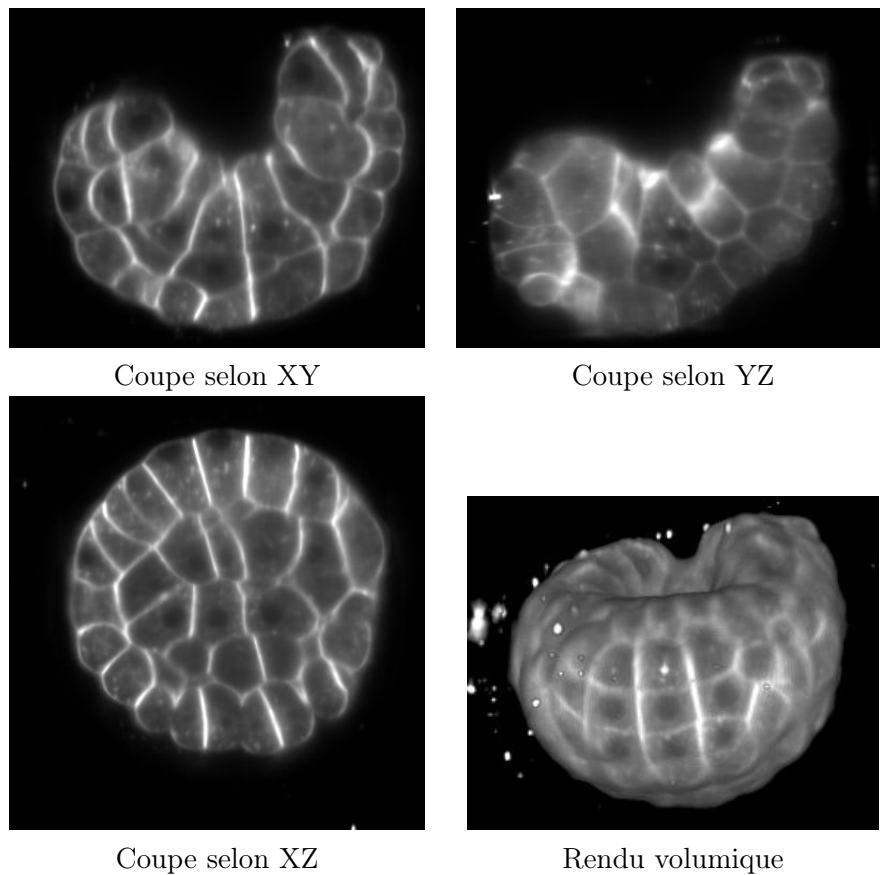


FIGURE 4.1 – Une image 3D d’embryon d’ascidie vue en coupe selon les plans XY, YZ et XZ. On observe que le signal des membranes cellulaires est plus fort pour les structures orthogonales à l’axe des X (anisotropie du processus d’acquisition). Les membranes de cellules à l’interface avec l’extérieur de l’embryon sont également plus faiblement contrastées que les membranes à l’interface entre deux cellules pour une même orientation spatiale. De plus, l’internalisation du marqueur dans le cytoplasme à l’intérieur des cellules provoque une dégradation du rapport signal-sur-bruit au cours du temps.

tres caractérisent formellement les structures planes — telles que les membranes que l'on cherche à rehausser — dont la particularité est d'être des structures bi-dimensionnelles, en se basant sur des matrices hessiennes [Mosaliganti 2010] ou des structures tensorielles [Drblíková 2009, Pop 2011].

Les premiers filtres basés sur la matrice hessienne (c'est-à-dire la matrice des dérivées secondes de l'intensité de l'image) ont été à l'origine proposés pour la détection de réseaux vasculaires en imagerie médicale [Lorenz 1997] en raison de leur capacité à caractériser des structures linéiques. L'exploitation de la matrice hessienne calculée en un point de l'image permet de construire une fonction de réponse linéique ad-hoc basée sur les valeurs propres de la matrice hessienne [Frangi 1998], ou encore d'intégrer un flux de diffusion (correspondant au vecteur gradient) autour du vecteur propre associé à la valeur propre la plus forte de la matrice hessienne pour la détection de lignes centrales de vaisseaux [Krissian 2000].

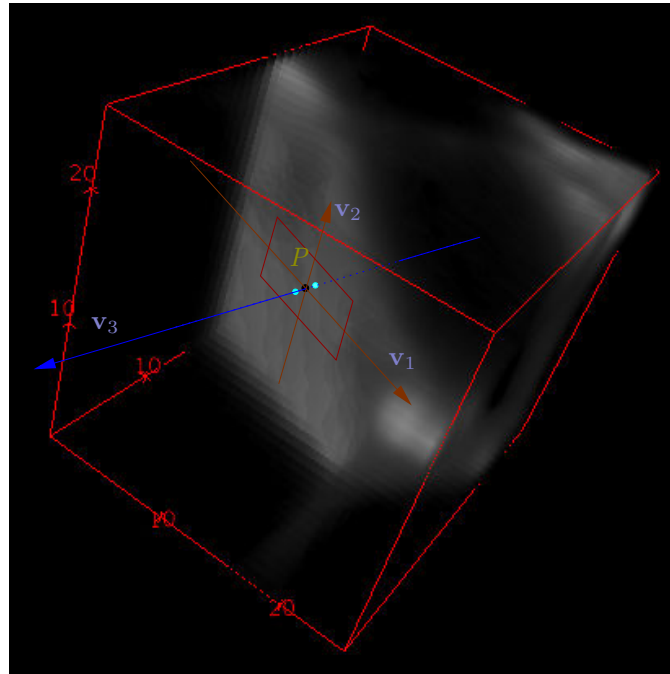
Pour la détection de structures planaires, [Mosaliganti 2010] propose une fonction de planéité directement adaptée de la formulation de [Frangi 1998]. Considérant une image $I : \Omega \rightarrow \mathbb{R}_+$, la matrice des dérivées secondes $H = H(P)$ au point $P \in \Omega$ est définie par :

$$H = \begin{pmatrix} \frac{\partial^2 I}{\partial x^2}(P) & \frac{\partial^2 I}{\partial x \partial y}(P) & \frac{\partial^2 I}{\partial x \partial z}(P) \\ \frac{\partial^2 I}{\partial x \partial y}(P) & \frac{\partial^2 I}{\partial y^2}(P) & \frac{\partial^2 I}{\partial y \partial z}(P) \\ \frac{\partial^2 I}{\partial x \partial z}(P) & \frac{\partial^2 I}{\partial y \partial z}(P) & \frac{\partial^2 I}{\partial z^2}(P) \end{pmatrix}. \quad (4.1)$$

La décomposition de H en valeurs propres $\{\lambda_1, \lambda_2, \lambda_3\}$ triées par valeur absolue croissante ($|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$) et vecteurs propres $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ fournit une interprétation locale de la variation du profil d'intensité de l'image au point P dans les directions portées par chacun des vecteurs propres de H (figure 4.2). Dans le cas où les membranes apparaissent comme des structures intenses sur fond sombre, les structures planaires recherchées sont caractérisées par une forte valeur propre négative λ_3 (le vecteur propre associé \mathbf{v}_3 donne la direction normale à la membrane) et par deux valeurs propres λ_1 et λ_2 proches de 0 (la membrane est localement tangente au plan $(P, \mathbf{v}_1, \mathbf{v}_2)$) — une structure linéique (un vaisseau) serait quant à elle caractérisée par deux fortes valeurs propres négatives et une valeur propre faible. La fonction de planéité de [Mosaliganti 2010] est donnée par :

$$R_{\text{Mos}}(P) = \begin{cases} 0 & \text{si } \lambda_3 \geq 0, \\ \left(1 - \exp\left(-\frac{\sum \lambda_i^2}{2\gamma^2}\right)\right) \exp\left(-\frac{|\frac{\lambda_2}{\lambda_3}|^2}{2\alpha^2}\right) \exp\left(-\frac{\left(\frac{\sqrt{|\lambda_1 \lambda_2|}}{|\lambda_3|}\right)^2}{2\beta^2}\right) \exp\left(-\frac{2c^2}{\lambda_3^2}\right) & \text{sinon.} \end{cases} \quad (4.2)$$

Dans la suite de cette section, nous adaptions l'approche qu'utilise [Krissian 2000] pour la détection de vaisseaux pour proposer une fonction de planéité alternative à celle donnée par l'équation 4.2.



(a) Membrane cellulaire dans une image 3D.

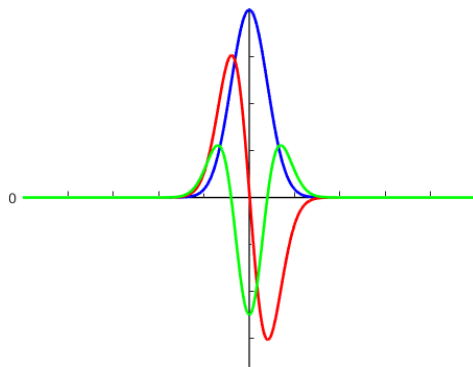
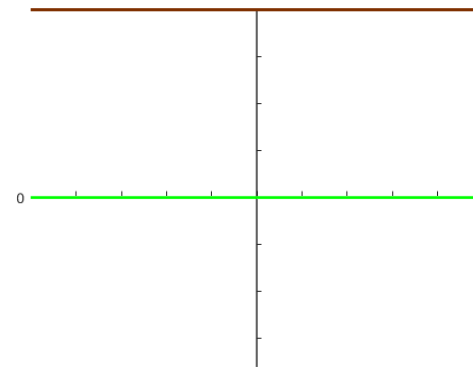
(b) Intensité selon la direction \mathbf{v}_3 (courbe bleue) et ses dérivées directionnelles du 1er ordre (courbe rouge) et du second ordre (courbe verte).(c) Intensité selon une direction du plan tangent à la membrane en P (courbe brune) et dérivées directionnelles du 1er ordre et du second ordre toutes deux nulles (courbe verte).

FIGURE 4.2 – Profils d'intensité et leurs dérivées directionnelles correspondantes dans les directions $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ correspondant aux vecteurs propres de la matrice hessienne H au point P .

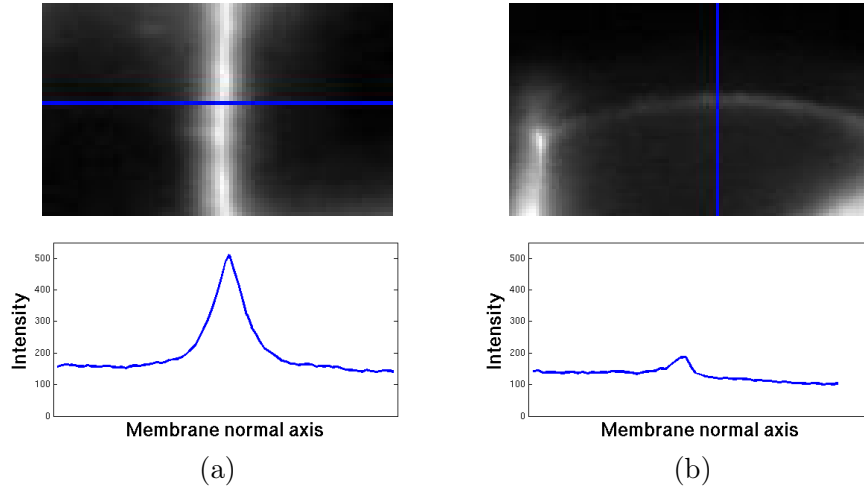


FIGURE 4.3 – Profils d'intensité dans la direction normale à des membranes. (a) Membrane séparant deux cellules. (b) Membrane séparant une cellule du fond.

4.2.2 Modèle planaire

La figure 4.3 présente des profils d'intensité typiques de membranes le long de la normale à une membrane. Ces profils représentatifs de signaux observés sont classiquement modélisés par une équation de type « gaussienne + constante ». Ce modèle est paramétré par un écart-type σ_0 , un coefficient multiplicatif α et un terme constant I_0 . L'équation modèle de variation d'intensité le long de l'axe normal d'une membrane centrée en 0 est donnée par :

$$I(x) = \frac{\alpha}{\sigma_0 \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma_0^2}\right) + I_0 = \alpha G_{\sigma_0}(x) + I_0, \quad (4.3)$$

où G_{σ_0} est le noyau gaussien d'écart type σ_0 .

4.2.3 Fonction de réponse

4.2.3.1 Étude à échelle fixe σ

À une échelle donnée σ , nous calculons pour chaque voxel P de l'image le vecteur gradient $\nabla_\sigma I(P)$ et la matrice hessienne $H(P) = \nabla_\sigma^2 I(P)$ qui lui sont associés par convolution de l'image avec des dérivées d'ordre 1 et 2 d'un noyau gaussien d'écart-type σ . L'analyse du système de valeurs propres et vecteurs propres de la matrice hessienne fournit des informations sur l'orientation locale de la structure présente en P . Notons λ_i (respectivement \mathbf{v}_i) la $i^{\text{ème}}$ valeur propre (respectivement son vecteur propre associé) de $\nabla_\sigma^2 I(P)$ vérifiant $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$. Les membranes apparaissent comme des structures planaires intenses sur fond sombre. Un voxel P appartenant à une membrane se caractérise donc par un large λ_3 négatif en raison de la variation d'intensité dans la direction normale à la membrane, et par de petites valeurs de $|\lambda_1|$ et $|\lambda_2|$ en raison de l'homogénéité locale de l'intensité dans les directions tangentes à

la membrane. Le vecteur propre \mathbf{v}_3 donne alors la direction normale à la membrane en P .

Considérant un voxel P et le couple $(\lambda_3, \mathbf{v}_3)$ extrait de la matrice hessienne $\nabla_\sigma^2 I(P)$, nous proposons une fonction additionnant une mesure de contours à une distance r_σ de part et d'autre de P par rapport à \mathbf{v}_3 , soit pour les points $M = P \pm r_\sigma \mathbf{v}_3$. La mesure de contours n'est autre que la dérivée directionnelle en M par rapport à la direction \overrightarrow{PM} :

$$R_\sigma(P) = \begin{cases} 0 & \text{si } \lambda_3 \geq 0, \\ \frac{1}{2} (\nabla_\sigma I(P - r_\sigma \mathbf{v}_3) \cdot \mathbf{v}_3 - \nabla_\sigma I(P + r_\sigma \mathbf{v}_3) \cdot \mathbf{v}_3). \end{cases} \quad (4.4)$$

Le paramètre r_σ correspond à la distance de P de laquelle la réponse doit être calculée et il doit être choisi selon le modèle planaire. Cette réponse est une adaptation du travail proposé par [Krissian 2000] pour l'extraction des structures linéiques. La valeur du gradient $\nabla_\sigma I(P \pm r_\sigma \mathbf{v}_3)$ est obtenue par interpolation trilineaire du vecteur gradient des voxels avoisinants.

4.2.3.2 Normalisation de la fonction de réponse

La fonction de réponse définie par l'équation 4.4 permet de filtrer des structures planaires à une échelle donnée. Lorsqu'on augmente le paramètre de convolution σ , l'intensité du signal et ses dérivées sont diminuées. Il est pourtant intéressant d'envisager un filtre planaire multi-échelle, au sens où l'on pourrait comparer des réponses obtenues par l'équation 4.4 avec des échelles variables. Lindeberg [Lindeberg 1994] est le premier à introduire la notion de dérivées normalisées, ou de γ -dérivées, dans ce but. L'idée est d'exprimer l'image I dans un système de coordonnées sans dimension, de telle sorte que les dérivées et fonctions de réponse qui en résultent sont indépendantes d'un facteur d'agrandissement. Cette notion de γ -dérivée a été reprise par [Krissian 2000] pour proposer une fonction de réponse linéique multi-échelle. Considérant la fonction de réponse R_σ calculée à l'échelle σ , l'expression de la fonction de réponse *normalisée* R_σ^n est donnée par :

$$R_\sigma^n = \sigma R_\sigma \quad (4.5)$$

4.2.3.3 Réponse multi-échelle

Soit σ_{\max} l'échelle qui maximise la réponse normalisée d'un voxel P appartenant à un plan. [Krissian 2000] montre que les frontières du plan (au sens de la maximisation de la norme du vecteur gradient) sont à une distance $\sqrt{\sigma_0^2 + \sigma_{\max}^2}$ de P dans la direction normale au plan modélisé par l'équation 4.3. Il montre également que la relation entre le paramètre de distance $r_{\sigma_{\max}}$ de l'équation 4.4 doit vérifier :

$$r_{\sigma_{\max}} = \sqrt{3} \sigma_{\max} \quad (4.6)$$

et que la réponse normalisée est maximisée pour $\sigma = \sigma_{\max} = 0,5\sigma_0$.

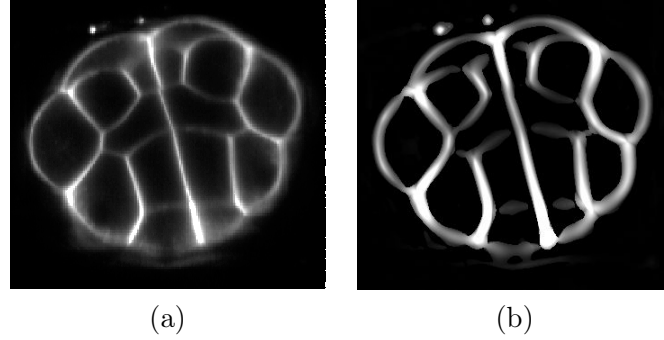


FIGURE 4.4 – Résultat de l’application de la fonction de planéité multi-échelle de l’équation 4.7. (a) Vue en coupe de l’image d’origine I . (b) Image filtrée.

On procède à une discrétisation logarithmique $\{\sigma_1, \dots, \sigma_h\}$ du domaine des échelles et pour chaque valeur de σ , on calcule la réponse normalisée R_σ^n de l’image. La réponse multi-échelle R^n au voxel P est donnée par :

$$R^n(P) = \max_{\sigma \in \{\sigma_1, \dots, \sigma_h\}} R_\sigma^n(P). \quad (4.7)$$

Le vecteur normal \mathbf{v}_3 associé au voxel P est le vecteur propre calculé à l’échelle σ_{\max} . La figure 4.4 montre un exemple de résultat d’application de la fonction de réponse donnée par l’équation 4.7.

4.3 Binarisation des membranes

La fonction de réponse proposée consiste en une somme de fonctions de réponse proportionnelles au gradient du profil d’intensité (c’est-à-dire proportionnelles à α dans l’équation-modèle 4.3). Les systèmes d’acquisition d’images en microscopie à fluorescence sont souvent anisotropes. Le profil d’intensité d’une membrane peut fortement varier dans une même image en fonction de l’orientation spatiale de la membrane (figure 4.5).

Nous proposons dans cette section d’extraire les plans centraux qui résultent de la fonction de planéité de la section 4.2 et de les binariser à l’aide d’un seuillage adaptatif qui prend en compte l’orientation spatiale de ces structures afin de rendre compte de l’anisotropie du signal induite par le système d’imagerie.

4.3.1 Extraction des plans centraux

La fonction de planéité définie par l’équation 4.7 est conçue de sorte que la réponse multi-échelle soit localement maximisée au niveau des voxels situés sur le plan central des membranes dans leur direction normale. À partir de la fonction de réponse multi-échelle R^n appliquée à l’image I , nous définissons la fonction E d’extraction des plans centraux de l’image donnée en un voxel P par :

$$E(P) = \begin{cases} 0 & \text{si } R^n(P) \leq R^n(P \pm \mathbf{v}_3), \\ R^n(P) & \text{sinon.} \end{cases} \quad (4.8)$$

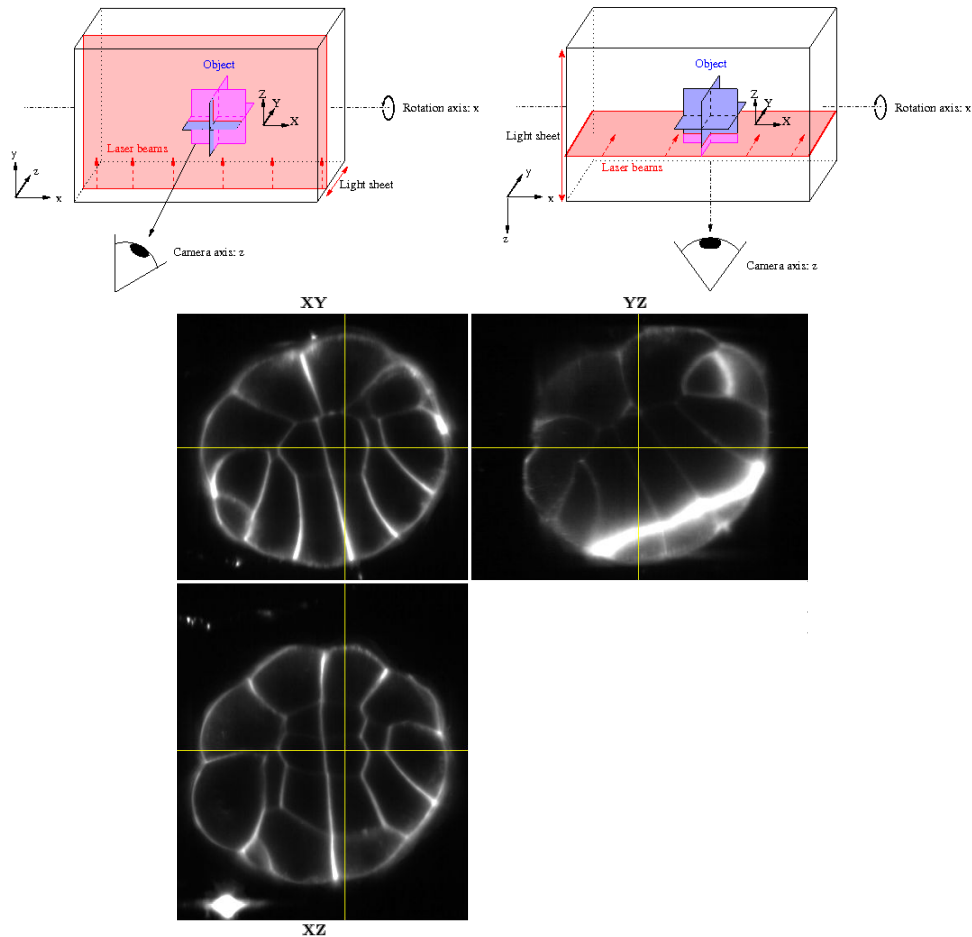


FIGURE 4.5 – Inhomogénéité de l'intensité des membranes d'une image due au système d'acquisition. Les schémas illustrent le fonctionnement du microscope MuViSPIM [Krzic 2012] qui image l'échantillon selon 4 orientations distinctes. Les orientations des prises de vue sont séparées d'un angle de 45° autour d'un axe fixe de l'image (ici l'axe des x). La rotation de la caméra autour de l'axe des x induit une illumination toujours intense des membranes normales à cet axe. À l'inverse, les membranes alignées avec l'axe de rotation de la caméra sont sensiblement moins illuminées selon les orientations de la caméra. Les images montrent les vues orthogonales d'une image d'embryon acquise par MuViSPIM après fusion des 4 vues [Krzic 2012]. L'intensité des membranes apparaît significativement plus grande pour les membranes normales à l'axe des x .

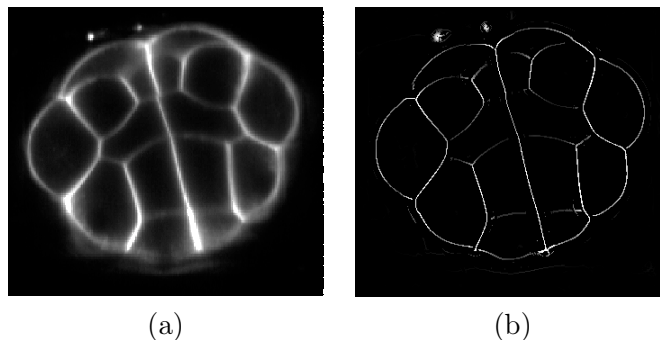


FIGURE 4.6 – Résultat de l'extraction des plans centraux donnée par l'équation 4.8. (a) Vue en coupe de l'image d'origine I . (b) Résultat de l'extraction des plans centraux de l'image filtrée.

Le vecteur \mathbf{v}_3 est le vecteur normal à la membrane au voxel P défini dans la section 4.2.3.3. Une illustration du résultat produit par la fonction d'extraction des plans centraux est donné par la figure 4.6.

4.3.2 Inhomogénéité de l'intensité des membranes

Dans le cadre de la segmentation des membranes, la binarisation des plans centraux peut être effectuée par un seuillage simple [Martinez-Sanchez 2013]. Cependant, l'intensité du signal d'une membrane est fortement liée à son orientation spatiale (figure 4.5). Dans un travail préliminaire [Michelin 2014], nous avons proposé de binariser l'image des plans centraux donnée par l'équation 4.8 en réalisant un seuillage par hystérésis, c'est-à-dire en sélectionnant deux seuils t_{\min} et t_{\max} tels qu'à un voxel P de l'image binarisée B_{hyst} est affectée une valeur de 1 si et seulement si $E(P) \geq t_{\max}$ ou $E(P) \geq t_{\min}$ et P est adjacent à un voxel de valeur 1 (les autres voxels sont affectés de la valeur 0). Cette méthode ne rend cependant pas compte de l'inhomogénéité de l'intensité induite par l'orientation spatiale des membranes. De plus, la sélection des seuils est réalisée manuellement. Nous présentons dans la suite de la section une méthode alternative de seuillage adaptatif selon l'orientation des structures avec calcul automatique des seuils.

4.3.3 Fonction de seuillage adaptatif

Afin de prendre en considération l'orientation spatiale des structures que l'on souhaite seuiller, nous définissons une fonction de seuillage adaptatif T qui prend en argument un vecteur normalisé $\mathbf{v} = [v_x, v_y, v_z]$ ($v_x^2 + v_y^2 + v_z^2 = 1$). La fonction de seuillage adaptatif T doit vérifier $T(\mathbf{v}) = T(-\mathbf{v})$, ainsi que les propriétés $T([1, 0, 0]) = t_x$, $T([0, 1, 0]) = t_y$ et $T([0, 0, 1]) = t_z$ — où les paramètres t_x , t_y et t_z sont des seuils « axiaux » à déterminer — afin que le seuillage d'une membrane orientée orthogonalement à l'un des axes du microscope ne dépende que du seuil axial correspondant à celui-ci. Une fonction qui découle naturellement des propriétés

sus-mentionnées est la fonction effectuant une combinaison linéaire des trois seuils axiaux selon l'expression :

$$T([v_x, v_y, v_z]) = t_x v_x^2 + t_y v_y^2 + t_z v_z^2. \quad (4.9)$$

Cette combinaison linéaire assure une somme des pondérations des seuils axiaux égale à 1. L'image binarisée B est ainsi définie par :

$$B(P) = \begin{cases} 1 & \text{si } E(P) \geq T(\mathbf{v}_3(P)), \\ 0 & \text{sinon.} \end{cases} \quad (4.10)$$

Par exemple, si un voxel P est orienté (au sens de $\mathbf{v}_3(P)$) selon l'axe des x , alors le seuil adaptatif calculé en ce point ne dépend que de t_x . Si dans l'image, les membranes orientées normalement à l'axe des x sont plus contrastées que les autres membranes comme sur la figure 4.5, alors les seuils axiaux devraient vérifier $t_x > t_{\{y,z\}}$. Dans ce qui suit, nous proposons une méthode de calcul des seuils axiaux adaptée développée pour répondre au problème de binarisation des membranes dans des images d'embryons acquises avec un microscope MuViSPIM [Krzic 2012].

4.3.4 Histogrammes axiaux

Considérant l'image E des plans centraux obtenue par l'équation 4.8, notons l'image des vecteurs normaux associés $\mathbf{V} = \{\mathbf{v}(P)\}_P$ telle que pour chaque voxel P , $\mathbf{v}(P) = \mathbf{v}_3(P)$.

Nous proposons de décomposer l'histogramme H de l'image scalaire E comme étant la somme d'*histogrammes axiaux* H_x , H_y et H_z en prenant en considération l'orientation spatiale de chaque voxel (algorithme 1). Les histogrammes axiaux peuvent être interprétés comme les contributions axiales de H . Considérant un voxel P , si P est orienté dans une direction privilégiée au sens de $\mathbf{v}(P)$, la contribution de $E(P)$ sera plus importante dans cette direction que dans les autres. La figure 4.7 présente le résultat de la décomposition de H en histogrammes axiaux sur une image d'embryon imagée par un microscope MuViSPIM [Krzic 2012]. Ce résultat met en évidence le comportement variable de E en fonction de l'orientation spatiale des membranes.

Algorithme 1 Construction des histogrammes axiaux H_x , H_y et H_z .

Init $H_x, H_y, H_z \leftarrow 0$
pour chaque voxel P t.q. $E(P) \neq 0$ **faire**
 $[v_x, v_y, v_z] := \mathbf{v}(P)$
 $r := E(P)$
 $H_x(r) \leftarrow H_x(r) + v_x^2$
 $H_y(r) \leftarrow H_y(r) + v_y^2$
 $H_z(r) \leftarrow H_z(r) + v_z^2$

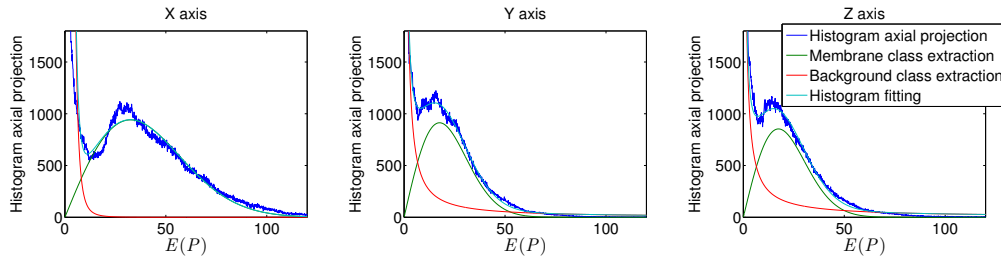


FIGURE 4.7 – Histogrammes axiaux (bleu) de l’image des plans centraux (équation 4.8) sur une image test d’embryon de *Phallusia mammillata* à un stade de 162 cellules. Les axes de projection y et z présentent un comportement similaire tandis que l’axe de projection des x présente un histogramme axial H_x plus large pour de grandes valeurs de $E(P)$, ce qui traduit des réponses de membranes plus fortes lorsque leur orientation spatiale est normale à l’axe des x . En rouge (resp. vert) : ajustement de la classe des plans centraux de fond (resp. de membrane). En cyan, la somme des courbes verte et rouge ajustant l’histogramme axial.

4.3.5 Seuils axiaux

Nous proposons une méthode automatique pour le calcul des seuils axiaux basée sur l’analyse des histogrammes axiaux calculés en section 4.3.4. Les voxels non-nuls de l’image E des plans centraux peuvent être classés dans deux groupes : ceux provenant de « vrais » signaux de membranes et ceux provenant de bruits ou du fond.

Nous calculons un ajustement de chaque histogramme axial comme étant la somme de deux fonctions. Nous postulons que l’une des deux fonctions modélise les plans centraux parasites, tandis que l’autre fonction modélise les plans centraux de « vraies » membranes. La figure 4.7 présente un exemple d’histogrammes axiaux caractéristique de l’ensemble des données d’embryons — obtenues par la technique d’imagerie MuViSPIM [Krzic 2012] — que nous avons eu à traiter. La forme des histogrammes axiaux suggère que la classe parasite peut être modélisée par une fonction de type $u \mapsto A/u^b$, et la classe des « vraies » membranes par une distribution de Rayleigh.

Après l’ajustement des histogrammes axiaux, à partir du moment où chacune des classes est modélisée par une fonction analytique déterminée, les seuils axiaux $t_{\{x,y,z\}}$ peuvent être calculés selon un critère objectif (rapport faux-positifs/faux-négatifs, pourcentage de faux-négatifs, ...). Dans notre application, nous privilégions un critère de type pourcentage de « vraies » membranes binarisées. Ce critère fait apparaître un unique paramètre p à fixer qui correspond au ratio de voxels provenant de vraies membranes que l’utilisateur espère binariser. La figure 4.8 présente le résultat de la binarisation des structures planaires sur un exemple d’image où le paramètre p a été fixé à 0,95.

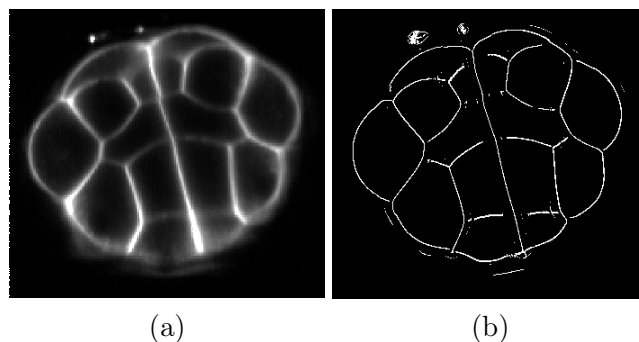


FIGURE 4.8 – Résultat de la binarisation des structures planaires donnée par l'équation 4.9. (a) Vue en coupe de l'image d'origine I . (b) Binarisation B des plans centraux par seuillage adaptatif.

4.4 Reconstruction d'une image d'intensité par vote de tenseurs

4.4.1 Introduction

L'organisation perceptuelle de l'information permet de regrouper les structures d'une image selon des indices de proximité et de bonne continuation. Les méthodes d'organisation (ou groupement) perceptuelle constituent une branche importante de la vision par ordinateur qui s'inspirent des théories gestaltistes [Koffka 1935] (psychologie de la forme). Ces théories descriptives définissent la notion de « champ perceptuel » analogue à celle des champs de force régissant les lois magnétiques et gravitationnelles. Ces « champs perceptuels » seraient à l'origine de l'impression de groupement spontané que donnent des éléments visuels selon leurs propriétés et leur agencement.

La nature des méthodes de groupement perceptuel consiste à regrouper des éléments isolés — appelés primitives ou *tokens* — d'une image suivant des modèles inspirés de critères perceptuels. Ces méthodes sont génériques, applicables à des données bruitées ou peu fiables et ont un champ d'application étendu. Il existe trois classes de méthodes de groupement perceptuel : les méthodes locales, semi-locales et globales. Une méthode locale cherchera à regrouper des paires de primitives en se basant uniquement sur les informations ne dépendant que d'elles [Heitger 1994]. Bien que rapide, ce type de méthode est cependant limité en terme de performance. Une méthode globale — comme la transformée de Hough généralisée — tient compte de l'ensemble des primitives pour effectuer les regroupements [Parent 1989]. Bien que performantes, la complexité algorithmique des méthodes globales limite cependant leur domaine d'applicabilité. Une méthode semi-locale permet de résoudre un problème global tout en émettant des hypothèses limitant les champs d'interaction entre primitives pour diminuer le coût algorithmique de la méthode.

Les méthodes semi-locales sont les plus répandues. [Ullman 1988] introduit pour la première fois la notion de « saillance » (*saliency*) structurelle dans le but de for-

mer des courbes saillantes à partir de primitives isolées. Dans [Parent 1989], les auteurs intègrent spatialement de l'information locale afin de déduire des structures linéiques et leur courbure. Ce travail est étendu à la déduction de surfaces en trois dimensions dans [Sander 1986]. [Sarkar 1994] présente l'idée de hiérarchisation organisationnelle des primitives, permettant d'extraire des structures variées, à l'aide d'un schéma de votes. La généralisation en 3D de cette méthode est cependant rendue difficile en raison de son coût algorithmique trop élevé dû à son incapacité à traiter simultanément les structures de type surfaces, courbes et jonctions. La méthode de votes de tenseurs (*tensor voting*) introduite par [Guy 1997] lève précisément cette limitation grâce à une représentation tensorielle des primitives qui encode simultanément les informations de surface, courbe et jonction. La technique du vote de tenseurs a évolué et été formalisée dans un certain nombre de travaux [Medioni 2000, Tong 2004, Mordohai 2006].

L'imagerie biomédicale 3D fait partie des domaines d'application du vote de tenseurs. Les applications principales de cet outil y sont la reconstruction de réseaux vasculaires [Risser 2008] et d'axones [Mottini 2012] pour les structures linéiques, et de membranes [Mosalganti 2012, Martinez-Sanchez 2014] pour les structures planaires. Dans cette section, nous nous intéressons à cette dernière application.

Le procédé de binarisation des membranes de la section 4.3 produit une image binaire B et une image vectorielle \mathbf{v} associée (où $\mathbf{v}(P) = \mathbf{v}_3(P)$ est défini en section 4.2.3.3) indiquant les orientations spatiales locales des structures planaires binarisées. Le but de cette section est de reconstruire une image en niveaux de gris à partir de B et \mathbf{v} à l'aide du processus de vote de tenseurs [Medioni 2000] qui comble les interruptions structurelles (discontinuités) de l'image binarisée tout en éliminant les détections de structures parasites. L'aspect computationnel de la méthode sera également abordé.

4.4.2 Les votes de tenseurs

4.4.2.1 Représentation structurelle

Dans le processus de vote de tenseurs, une image est encodée par une image de tenseurs d'ordre 2 (matrices 3×3 pour des images 3D, supposé être le cas ici) définis positifs \mathbf{T} . À un voxel P de l'image, $\mathbf{T}(P)$ décrit l'information structurelle contenue en P . Considérons le système de valeurs propres et vecteurs propres $(\kappa_i, \mathbf{e}_i)_{i \in \{1,2,3\}}$ de $\mathbf{T}(P)$ classés tels que $0 \leq \kappa_1 \leq \kappa_2 \leq \kappa_3$. Le tenseur $\mathbf{T}(P)$ peut être réécrit comme une combinaison linéaire de 3 tenseurs génériques appelés tenseur « balle » (*ball*), « plat » (*plate*) et « bâton » (*stick*) notés respectivement \mathbf{T}_B , \mathbf{T}_P et \mathbf{T}_S :

$$\mathbf{T}(P) = \underbrace{\kappa_1 \sum_{i \in \{1,2,3\}} \mathbf{e}_i \mathbf{e}_i^\top}_{\mathbf{T}_B} + (\kappa_2 - \kappa_1) \underbrace{\sum_{i \in \{2,3\}} \mathbf{e}_i \mathbf{e}_i^\top}_{\mathbf{T}_P} + (\kappa_3 - \kappa_2) \underbrace{\mathbf{e}_3 \mathbf{e}_3^\top}_{\mathbf{T}_S} \quad (4.11)$$

où dans la base $[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$,

$$\mathbf{T}_B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \mathbf{T}_P = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ et } \mathbf{T}_S = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Le tenseur balle \mathbf{T}_B encode une structure qui ne possède pas de direction préférentielle, ce qui correspond à une jonction. Le tenseur plat \mathbf{T}_P encode une structure qui se propage préférentiellement dans sa direction normale, ce qui correspond à une structure linéique de direction \mathbf{e}_1 . Le tenseur bâton \mathbf{T}_S encode une structure qui se propage préférentiellement dans le plan normal à son orientation, ce qui correspond à une structure plane de normale \mathbf{e}_3 . À un voxel P , les coefficients κ_1 , $(\kappa_2 - \kappa_1)$ et $(\kappa_3 - \kappa_2)$ représentent la proportion de chacune de ces structures encodées en P .

4.4.2.2 Initialisation de l'image de tenseurs

À l'origine du procédé de vote de tenseurs, la première étape consiste à initialiser l'image de tenseurs \mathbf{T}_{init} associée à l'image traitée. En fonction de l'information contenue en chaque voxel de l'image de départ, le tenseur initial $\mathbf{T}_{\text{init}}(P)$ peut être nul (absence d'information structurelle), orienté (tenseur plat ou bâton) ou non-orienté (balle) selon l'information dont on dispose.

Dans notre contexte, l'information provient d'une image binaire B associée à une image vectorielle \mathbf{v} qui renseigne de la normale calculée en chaque point binarisé de B . L'information qu'on souhaite encoder correspond à des structures planes. Nous construisons donc l'image de tenseurs initiale \mathbf{T}_{init} :

$$\mathbf{T}_{\text{init}}(P) = B(P)\mathbf{v}(P)\mathbf{v}(P)^\top, \quad (4.12)$$

de sorte qu'un voxel P binarisé de B (c'est-à-dire $B(P) = 1$) soit encodé par un tenseur bâton — structure plane — de valeurs propres $(\kappa_1, \kappa_2, \kappa_3) = (0, 0, 1)$ avec $\mathbf{e}_3 = \mathbf{v}(P)$.

4.4.2.3 Champs de tenseurs votants

Après l'initialisation de l'image de tenseurs, la méthode de votes de tenseurs proposée par [Medioni 2000] consiste en une étape de propagation de l'information locale dans les voisinages : chacun des voxels encodés par un tenseur non-nul émet un champ de tenseurs votants.

Le champ de vote émis par un tenseur permet la propagation de l'information dans la (les) direction(s) préférentielle(s) de la structure encodée par le tenseur. Un paramètre d'échelle σ_{TV} régule la taille du champ de vote, définissant la notion de voisinage d'une primitive.

Dans notre cadre, l'image initiale des tenseurs votants contient exclusivement des tenseurs nuls ou bâtons. Nous nous limitons donc à la définition donnée par [Medioni 2000] d'un champ de tenseurs issu d'un tenseur bâton. Le vote reçu au

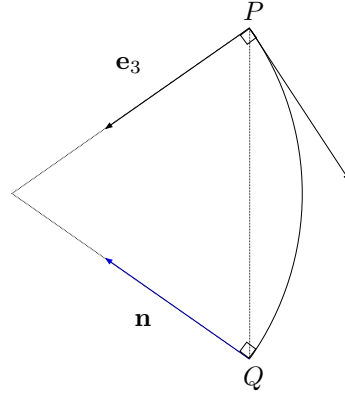


FIGURE 4.9 – Construction géométrique de l'arc reliant P au point Q et du vecteur \mathbf{n} .

point Q provenant du point P encodé par le tenseur bâton $\mathbf{T}_{\text{init}}(P) = \mathbf{T}_S = \mathbf{e}_3 \mathbf{e}_3^\top$ (équation 4.11) est donc un tenseur bâton décrivant une surface passant par P et Q et de normale \mathbf{e}_3 en P . On choisit une surface de courbure minimale, donc définie par un arc de cercle joignant P et Q . Le vote est pondéré, et décroît avec la distance géodésique ainsi que la courbure :

$$\mathbf{S}_{\sigma_{\text{TV}}}(\overrightarrow{PQ}, \mathbf{T}_S) = B(P) \exp - \frac{s^2 + c\kappa^2}{\sigma_{\text{TV}}} \mathbf{nn}^\top, \quad (4.13)$$

où s désigne la longueur d'arc et κ désigne la courbure de l'arc de cercle qui joint P à Q et qui est normal à \mathbf{e}_3 (équations 4.14 et 4.15). \mathbf{n} est le vecteur unité orthogonal à l'arc de cercle en Q (équation 4.16, voir figure 4.9). Le paramètre c contrôle le rapport d'influence de la proximité sur la courbure et est fixé à $c = -16 \log(0,1) \times (\sigma_{\text{TV}} - 1) \times \pi^{-2}$ [Medioni 2000]. Notons qu'aucun vote n'est reçu par Q si l'angle θ entre \mathbf{e}_3 et \mathbf{n} est supérieur à 90° :

$$s = \begin{cases} \frac{\frac{\theta}{2} \|\overrightarrow{PQ}\|}{\sin(\frac{\theta}{2})} & \text{si } \theta \neq 0 \\ \|\overrightarrow{PQ}\| & \text{sinon} \end{cases}. \quad (4.14)$$

$$\kappa = \frac{2 \sin(\frac{\theta}{2})}{\|\overrightarrow{PQ}\|}. \quad (4.15)$$

$$\mathbf{n} = \left(\mathbf{I}_3 - 2 \frac{\overrightarrow{PQ} \overrightarrow{PQ}^\top}{\|\overrightarrow{PQ}\|^2} \right) \mathbf{e}_3. \quad (4.16)$$

On observe sur la figure 4.10 le champ de tenseurs votants généré par un tenseur bâton. Ces champs de tenseurs s'agrègent par simple addition sur une carte de cumul des votes \mathbf{T}_{res} dense, où chaque voxel Q de l'image est encodé par un tenseur d'expression :

$$\mathbf{T}_{\text{res}}(Q) = \sum_P \mathbf{S}_{\sigma_{\text{TV}}}(\overrightarrow{PQ}, \mathbf{T}_{\text{init}}(P)). \quad (4.17)$$

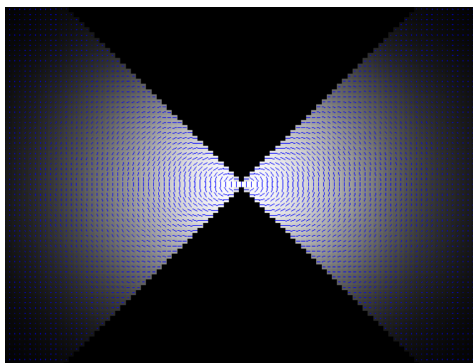


FIGURE 4.10 – Champ de vote généré par un tenseur bâton (vu en coupe). Les niveaux de gris donnent l’amplitude du vote et les bâtonnets bleus indiquent l’orientation des votes reçus.

4.4.2.4 Interprétation de la carte de cumul des votes

Chaque voxel de \mathbf{T}_{res} est décomposé en valeurs propres selon l’équation 4.11 afin de calculer l’adéquation de la structure à une jonction (composante *ball*), une courbe (composante *plate*) ou une surface (composante *stick*). On peut donc construire trois nouvelles images de « saillance » \mathcal{J} , \mathcal{C} et \mathcal{S} exprimant les propriétés structurales locales de l’image tenseur :

$$\mathcal{J}(P) = \kappa_1 (\mathbf{T}_{\text{res}}(P)), \quad (4.18)$$

$$\mathcal{C}(P) = \kappa_2 (\mathbf{T}_{\text{res}}(P)) - \kappa_1 (\mathbf{T}_{\text{res}}(P)), \quad (4.19)$$

$$\mathcal{S}(P) = \kappa_3 (\mathbf{T}_{\text{res}}(P)) - \kappa_2 (\mathbf{T}_{\text{res}}(P)). \quad (4.20)$$

Pour la reconstruction de membranes, les outils développés dans la littérature exploitent directement l’image de saillance surfacique \mathcal{S} de l’équation 4.20 [Mosaliganti 2012, Martinez-Sanchez 2014]. Dans un premier temps, nous avons également exploité l’image \mathcal{S} pour la reconstruction d’une image d’intensité. Cependant, cette image échoue à reconstruire correctement les régions situées à l’interface entre plus de deux cellules car l’organisation de l’information correspond localement à une intersection de plans (figure 4.11b). Pour extraire à la fois les interfaces entre deux cellules (localement planes) et les interfaces entre davantage de cellules (orientation préférentielle dans la direction de l’intersection entre les plans), nous proposons de définir une image de « saillance membranaire » \mathcal{M} calculée comme étant la somme des images de saillance surfacique et linéique (figure 4.11c) :

$$\mathcal{M}(P) = \mathcal{S}(P) + \mathcal{C}(P). \quad (4.21)$$

L’inconvénient de l’extracteur \mathcal{M} est qu’il peut être amené à extraire des structures linéiques éventuellement présentes dans l’image d’origine en raison du terme \mathcal{C} . Cependant, dans le cadre de ce travail, nous traitons des images dans lesquelles de telles structures parasites sont rares. De plus, le filtrage de l’image d’origine

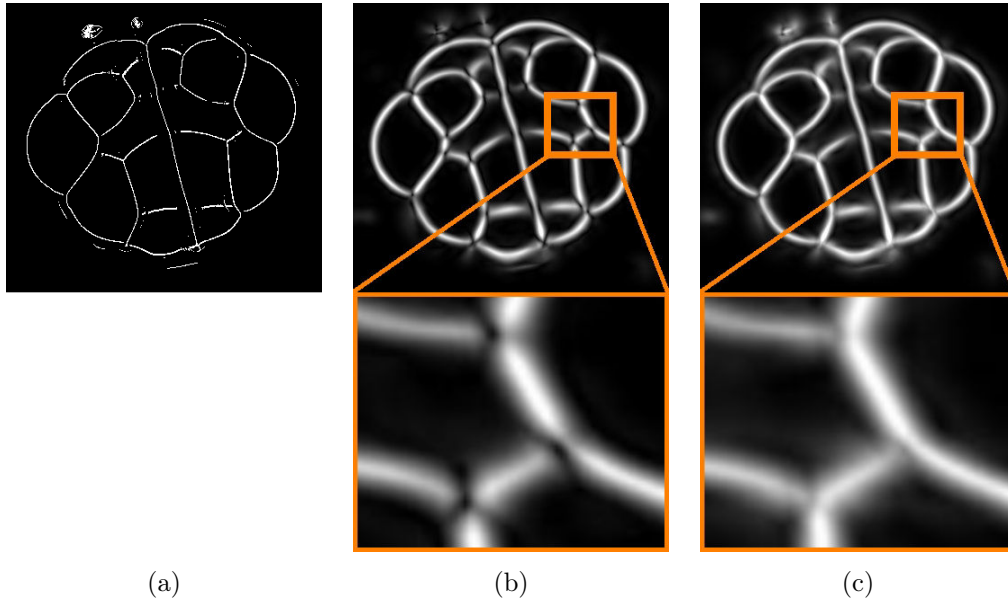


FIGURE 4.11 – Extracteur de membranes. (a) Image binaire B en entrée (vue en coupe). (b) Image de saillance surfacique \mathcal{S} (équation 4.20). (c) Image de saillance membranaire \mathcal{M} (équation 4.21). La reconstruction des structures planaires de l'image d'entrée est réalisée correctement par \mathcal{S} et \mathcal{M} . Toutefois, \mathcal{M} permet une bien meilleure reconstruction des jonctions entre plus de deux cellules que \mathcal{S} .

par la fonction de planéité définie par l'équation 4.7 est réalisé en amont du vote de tenseurs, ce qui permet d'éliminer ces éventuelles structures. La comparaison visuelle des résultats de reconstruction obtenus avec de extracteurs \mathcal{S} et \mathcal{M} (figure 4.11) montre l'effet positif de l'ajout de terme de saillance linéique sur l'extracteur de membranes.

4.4.2.5 Coût algorithmique du vote de tenseurs

Le vote de tenseurs se décompose en trois étapes : (1) l'initialisation de l'image des tenseurs \mathbf{T}_{init} ; (2) l'étape d'accumulation des votes de tenseurs ; (3) la conversion de la carte de cumul des votes \mathbf{T}_{res} en une image scalaire de « saillance membranaire » \mathcal{M} . Les étapes (1) et (3) sont incompressibles et ont un coût linéairement dépendant du nombre de voxels qui composent l'image de départ. L'étape (2) concentre toutes les difficultés d'implémentation et de coût de calcul de la méthode.

Processus d'accumulation des votes. Plusieurs approches peuvent être envisagées pour l'implémentation de l'étape de vote des tenseurs. Des recherches visent à trouver une solution analytique pour le calcul de \mathbf{T}_{res} en fonction de \mathbf{T}_{init} [Wu 2012], qui présenterait de nombreux avantages : rapidité de calcul, solution exacte ou encore applicabilité au calcul différentiel. Cependant, l'expression analytique proposée dans [Wu 2012] était erronée [Maggiori 2014] et à ce jour, aucune solution analytique

valide n'a été déterminée pour le problème du vote de tenseurs.

L'approche intuitive consiste donc à calculer systématiquement le champ de vote émis par chaque tenseur non-nul de \mathbf{T}_{init} pour l'ajouter à la carte de cumul des votes \mathbf{T}_{res} . Le coût du calcul des champs de vote est cependant bien plus grand que le coût que représente l'addition de ces champs à \mathbf{T}_{res} . La solution alternative que nous utilisons consiste à pré-calculer des champs de tenseurs et à les stocker dans une table de correspondance : on discrétise les directions de l'espace et on fait correspondre à chacune de ces directions un champ de tenseurs émis par un tenseur bâton porté par cette direction. Le champ de vote émis par un tenseur de \mathbf{T}_{init} est alors approché par le champ de vote de la table de correspondance dont l'orientation est la plus proche [Mordohai 2006]. Le coût constant du pré-calcul des champs de vote devient ainsi très petit en comparaison du coût de l'addition des votes qui est proportionnel au nombre de tenseurs non-nuls (c'est-à-dire votants) de \mathbf{T}_{init} . La complexité en temps de calcul du vote de tenseurs est donc — à une constante près — linéairement dépendant du nombre de primitives contenues dans l'image binaire B servant à l'initialisation de \mathbf{T}_{init} dans l'équation 4.12.

Échantillonnage des primitives pour l'optimisation du coût computationnel. Dans la méthode ACME proposée par [Mosaliganti 2012], l'étape de vote de tenseurs est réalisée avec une image de tenseurs \mathbf{T}_{init} initialisée directement par le résultat du filtrage R_{Mos} donné par l'équation 4.2. Cette initialisation dense fait donc intervenir un grand nombre de primitives (environ la moitié des voxels de l'image initiale) et génère des votes provenant d'informations parasites.

L'une des contributions de notre travail par rapport à [Mosaliganti 2012] est de proposer une sélection plus fine des primitives. Nous éliminons dans un premier temps la majorité de l'information parasite grâce aux étapes d'extraction et de binarisation des plans centraux de la fonction planéité (section 4.3). L'utilisation de l'image binarisée B pour l'initialisation du vote de tenseurs réduit d'un facteur environ égal à 50 le nombre de primitives et donc la complexité du problème en comparaison de la méthode ACME, tout en satisfaisant mieux l'esprit original de la méthode de vote de tenseurs qui suppose de reconstruire une carte dense provenant d'une information éparse [Medioni 2000]. Cependant, le vote de tenseurs demeure l'étape la longue de notre chaîne de traitement avec un temps de calcul de l'ordre de l'heure pour des images de taille importante (de l'ordre de $200^3 \approx 10^7$ voxels par image) avec une échelle de vote $\sigma_{\text{TV}} = 15$ (choisi en fonction de la taille des discontinuités à reconstruire dans l'image).

Pour diminuer encore le coût de l'étape de vote de tenseurs, nous avons étudié l'influence de l'échantillonnage des primitives provenant de l'image binaire B (équation 4.10) pour l'étape d'initialisation sur la qualité de reconstruction des membranes. Le niveau d'échantillonnage idéal de B dépend de l'échelle de vote σ_{TV} . La condition nécessaire à la bonne reconstruction d'une image \mathcal{M} est que les interruptions de signal engendrées par l'échantillonnage de B doivent avoir une taille plus petite que le paramètre σ_{TV} (figure 4.12). Cette condition assure que le signal de

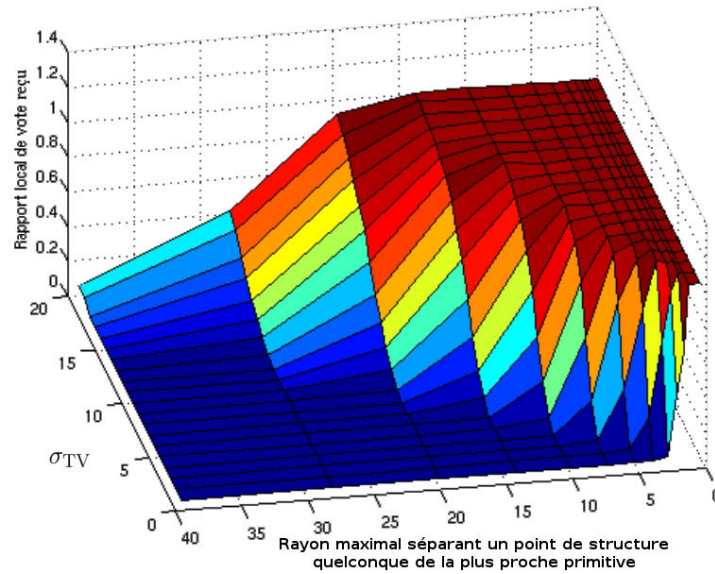


FIGURE 4.12 – Simulation de l'évolution du rapport des mesures de saillance surfacique \mathcal{S} reçues par un point quelconque placé sur une structure plane et par un point de primitive en fonction de l'échelle de vote σ_{TV} et de la distance maximale séparant un point du plan de la primitive (i.e. point votant) la plus proche. Le rapport de vote reçu est voisin de 1 lorsque cette distance est inférieure à σ_{TV} et décroît rapidement au delà.

l'image reconstruite en un point traversé par un plan ne dépende pas de la distance entre le point et la plus proche primitive.

L'échantillonnage de B peut être réalisé par tirage aléatoire (loi uniforme) ou régularisé [Cachia 2003] des primitives (figure 4.13). Le ratio d'échantillonnage maximal dépend de l'échelle de vote σ_{TV} . Le gain en temps de calcul du vote de tenseurs peut par exemple atteindre un facteur 30 pour $\sigma_{TV} = 15$ (paramètre utilisé dans nos expériences) tout en observant les mêmes qualités de reconstruction.

4.5 Segmentation des cellules

Une fois la procédure de reconstruction d'une image d'intensité réalisée (section 4.4), nous appliquons l'algorithme classique de la ligne de partage des eaux marquée [Adams 1994] sur l'image \mathcal{M} donnée par l'équation 4.21. La segmentation se décompose en deux parties : (1) on extrait les marquages (graines) qui seront utilisés pour l'initialisation de la procédure ; (2) on procède à la segmentation de l'image par la procédure décrite dans [Adams 1994]. Notre procédure est largement inspirée de l'algorithme MARS décrit dans [Fernandez 2010a] pour la segmentation de cellules de méristèmes et réutilisé par [Guignard 2015] pour la segmentation de cellules d'embryons d'ascidies.

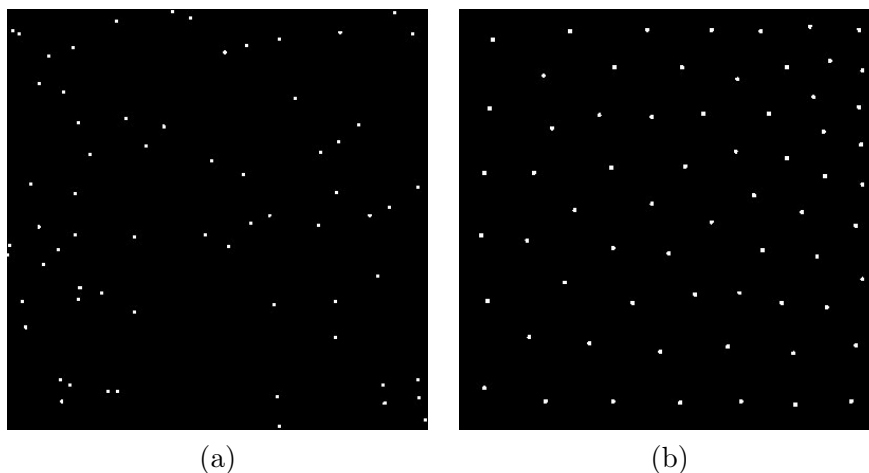


FIGURE 4.13 – Exemple de l'échantillonnage d'un carré. (a) Échantillonnage aléatoire. (b) Échantillonnage régulier [Cachia 2003].

4.5.1 Détection de graines

La segmentation de \mathcal{M} débute par la détermination de graines qui seront à l'origine de chacune des régions segmentées. Nous appliquons l'opérateur morphologique h -minima [Soille 1999] qui permet d'extraire de l'image \mathcal{M} , préalablement légèrement lissée par un filtrage gaussien, les minima régionaux dont le contraste est supérieur à un paramètre de seuil h_{\min} . L'étape de segmentation des membranes (section 4.3) qui précède le vote de tenseurs peut parfois manquer la détection des membranes les moins contrastées de l'image d'origine. Pour prévenir l'éventualité d'une « ouverture » dans l'image \mathcal{M} reliant deux cellules voisines, nous calculons une carte des distances dans l'image des h -minima, puis nous calculons et étiquetons par des labels les h -maxima de la carte des distances, ce qui détermine les graines utilisées en entrée de l'algorithme de ligne de partage des eaux.

4.5.2 Ligne de partage des eaux

La chaîne de traitement que nous proposons pour la segmentation de cellules de tissus épithéliaux s'achève sur l'étape de segmentation par la méthode de ligne de partage des eaux (figure 4.14). Celle-ci est réalisée sur l'image utilisée pour la détection des graines. Nous utilisons un algorithme par immersion [Vincent 1991] pour cette dernière étape.

4.6 Expériences et résultats

4.6.1 Données

Cinq embryons de *Phallusia mammillata* ont été imagés à l'aide de la technique de microscopie MuViSPIM [Krzic 2012]. Le pas de temps entre deux acquisitions

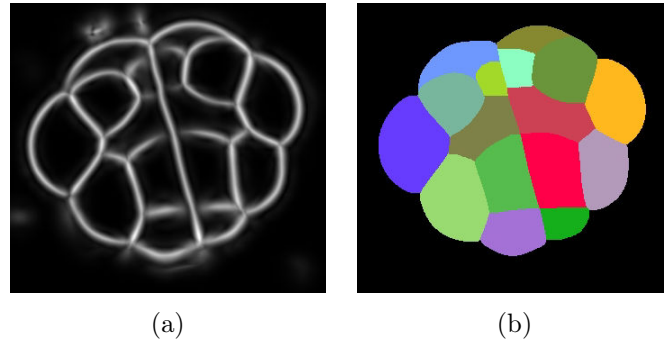


FIGURE 4.14 – Segmentation des cellules. (a) Image de saillance membranaire \mathcal{M} en entrée de procédure (vue en coupe). (b) Image segmentée.

successives est de l'ordre d'une minute. Pour chaque image, l'acquisition est réalisée selon quatre angles fournissant chacun un point de vue de taille $1200 \times 1200 \times 200$ voxels et de résolution $0,26\mu\text{m} \times 0,26\mu\text{m} \times 1\mu\text{m}$. Les points de vue sont ensuite fusionnés pour ne former qu'une seule image ayant une résolution isotrope et dont les défauts d'acquisition sont atténués. Les membranes cellulaires sont marquées par un colorant lipophile (FM4-64) pour le premier embryon, et par un marqueur génétique (PH-GFP) pour les autres. Au total, nous avons acquis 115 images :

- 1^{er} embryon : 99 images du stade 32 cellules au stade 172 cellules ;
- 2^{ème} embryon : 9 images du stade 64 cellules au stade 112 cellules ;
- 3^{ème} embryon : 5 images du stade 64 cellules au stade 70 cellules ;
- 4^{ème} embryon : 1 images au stade 114 cellules ;
- 5^{ème} embryon : 1 image au stade neurula (environ 450 cellules).

4.6.2 Évaluation de la binarisation des membranes

Parmi les défauts du marquage des membranes par un colorant lipophile, l'internalisation du colorant dans le cytoplasme des cellules dégrade le rapport signal-sur-bruit au cours du temps et des divisions cellulaires. Nous avons choisi d'évaluer notre méthode de segmentation des membranes sur des données provenant du premier embryon aux 1^{er} (32 cellules) et 91^e (162 cellules) instants de la série. Ces deux images sont représentatives de la diversité des conditions d'imagerie des embryons d'ascidies : la première image est acquise dans des conditions favorables, tandis que l'acquisition de la seconde image est dégradée par l'internalisation du marqueur lipophile dans le cytoplasme des cellules. Une segmentation experte (correction manuelle de la segmentation des cellules issue de la méthode MARS [Fernandez 2010a]) a été réalisée sur les deux images afin de posséder une vérité terrain (*ground truth*, abrégée en GT).

4.6.2.1 Méthodologie

La fonction de seuil adaptatif T proposée (équation 4.9) est une combinaison linéaire de trois seuils axiaux t_x , t_y et t_z fonction de l'orientation structurelle \mathbf{v} du voxel considéré. Si on note $t_{\text{low}} = \min_{w \in \{x,y,z\}} \{t_w\}$ et $t_{\text{high}} = \max_{w \in \{x,y,z\}} \{t_w\}$,

$$\forall \mathbf{v} \in \mathbb{R}^3 / \|\mathbf{v}\| = 1, \quad t_{\text{low}} \leq T(\mathbf{v}) \leq t_{\text{high}}. \quad (4.22)$$

Dans ce qui suit, B_{low} et B_{high} désignent les ensembles de voxels résultant du seuillage de l'image des plans centraux E (équation 4.8) par les seuils t_{low} et t_{high} respectivement. B_{adaptive} désigne l'ensemble obtenu par l'application du seuillage adaptatif donné par l'équation 4.10. La relation 4.22 garantit $B_{\text{high}} \subseteq B_{\text{adaptive}} \subseteq B_{\text{low}}$. B_{GT} est l'ensemble des points établis comme points de membranes grâce à la vérité terrain (GT).

Nous considérons que la binarisation la plus sélective B_{high} exclut la plupart des points de plans centraux parasites, mais aussi les points appartenant aux membranes les moins contrastées. La binarisation la moins sélective B_{low} conserve la plupart des plans centraux de membranes mais aussi des structures parasites. Notre postulat est que la binarisation par seuillage adaptatif B_{adaptive} que nous proposons permet (1) de conserver la plupart des plans centraux de membranes et (2) d'exclure la plupart des structures parasites provenant de $B_{\text{low}} \setminus B_{\text{high}}$.

Dans ce qui suit, nous utilisons une mesure de distance D entre un voxel binarisé P et la vérité terrain donnée par l'ensemble de points B_{GT} pour quantifier les erreurs de détection de membranes. En fixant un critère de distance d définissant la tolérance sur la localisation des membranes binarisées par rapport à la vérité terrain, nous décomposons un ensemble de points B en l'union de deux ensembles de points :

$$B = (B)_{\leq d} \cup (B)_{> d}, \quad (4.23)$$

où $(B)_{\leq d} = \{P \in B \mid D(P, B_{\text{GT}}) \leq d\}$ et $(B)_{> d} = \{P \in B \mid D(P, B_{\text{GT}}) > d\}$. L'ensemble $(B)_{\leq d}$ peut être interprété comme le sous-ensemble de B correspondant aux « bonnes » détections et l'ensemble $(B)_{> d}$ forme alors le sous-ensemble de B de « mauvaises » détections au sens de B_{GT} avec la tolérance d sur les erreurs de détection. En calculant l'histogramme de $D(B) = \{D(P, GT)\}_{P \in B}$, nous pouvons mesurer le rapport de « bonnes / mauvaises » détections de B .

Nous vérifions notre postulat en mesurant :

- (1) le taux de conservation des bonnes structures
- (2) le taux de suppression des mauvaises structures

de B_{adaptive} par rapport à $B_{\text{low}} \setminus B_{\text{high}}$, avec une tolérance d fixée à 4 voxels ($1\mu\text{m}$) qui traduit raisonnablement l'incertitude portant sur la précision de la localisation des membranes dans la vérité terrain utilisée.

4.6.2.2 Résultats

La figure 4.15 présente pour les deux images de test les histogrammes des distances de $B_{\text{low}} \setminus B_{\text{high}}$, $B_{\text{adaptive}} \setminus B_{\text{high}}$ et $B_{\text{low}} \setminus B_{\text{adaptive}}$ par rapport à B_{GT} . Ces

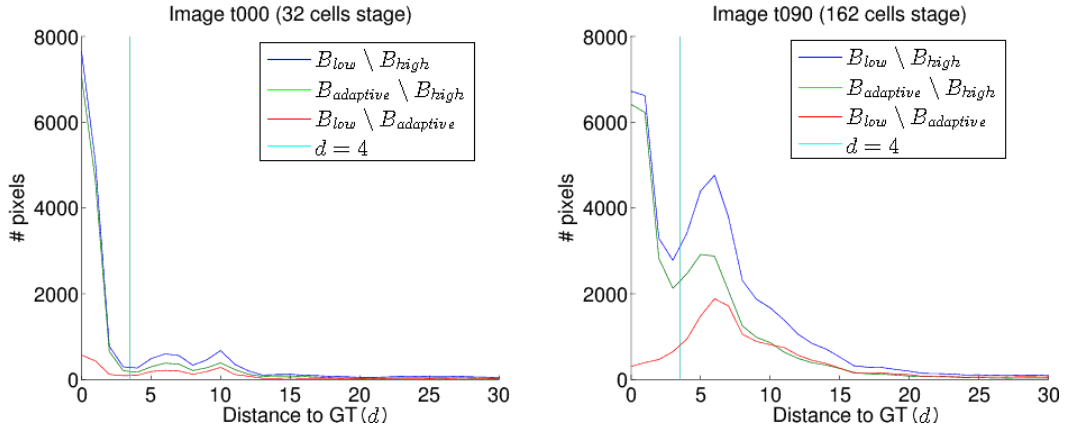


FIGURE 4.15 – Histogrammes des distances à la vérité terrain. À gauche : stade 32 cellules. À droite : stade 162 cellules. Histogrammes bleu, vert, rouge respectivement pour les ensembles $B_{low} \setminus B_{high}$, $B_{adaptive} \setminus B_{high}$, $B_{low} \setminus B_{adaptive}$. La ligne verticale délimite la tolérance sur les erreurs de détection.

histogrammes montrent que pour les deux images testées, le seuillage adaptatif conserve la plupart des « bonnes » structures présentes dans B_{low} tout en réduisant significativement les structures parasites dans la binarisation finale. Plus précisément, pour une tolérance de $1\mu\text{m}$ sur la détection des membranes, le taux de conservation des « bonnes » structures atteint 90,6% au stade 32 cellules et 87,9% au stade 162 cellules, alors que l'élimination des structures parasites atteint 38,9% au stade 32 cellules et 44,8% au stade 162 cellules. Ces résultats démontrent une amélioration significative de l'étape de binarisation des membranes apportée par la méthode de seuillage adaptatif proposée qui se base sur l'orientation des structures dans le repère de l'image et sur la décomposition axiale de l'histogramme de la carte des plans centraux.

4.6.3 Évaluation du pré-traitement

Une partie de nos contributions porte sur le pré-traitement de l'image d'origine précédant la reconstruction et la segmentation d'une image scalaire. Dans cette partie, nous évaluons les aspects comprenant la mise en place d'une nouvelle fonction de planéité R (équation 4.7) et le choix d'un nouveau modèle de sélection des primitives utilisées pour le vote de tenseurs par rapport au travail présenté dans [Mosaliganti 2012].

Pour évaluer l'apport de ces contributions, nous adoptons un plan de test (figure 4.16) qui consiste à comparer les résultats de segmentation obtenus avec les stratégies de pré-traitement suggérées avec ceux obtenus par le pré-traitement de [Mosaliganti 2012]. Nous comparons également ces résultats à ceux obtenus par la méthode de référence MARS [Fernandez 2010a] (méthode utilisée par [Guignard 2015] sur nos données de test). La comparaison des résultats de segmentation des cellules est opérée sur les deux images d'embryons présentées en

section 4.6.2.

4.6.3.1 Méthodologie

Pour chacun des flux de traitement du plan de test (figure 4.16), nous évaluons quantitativement la performance de la segmentation finale par rapport à la vérité terrain. Nous faisons porter l'évaluation sur la capacité de la chaîne de traitement à détecter correctement les cellules. À l'aide de la méthode de comparaison des segmentations présentée au chapitre 3, nous mesurons :

- les bonnes détections (*True Detection*, TD) : bijection entre cellule de vérité terrain et cellule segmentée ;
- les sur-détections (*Over Detection*, OD) : une cellule de vérité terrain est le siège de plusieurs détections. On comptabilise autant de sur-détections qu'il y a de segmentations en trop ;
- les sous-détections (*Under Detection*, UD) : deux cellules de vérité terrain sont segmentées en une seule région, ou une cellule est située dans le label de fond de la segmentation.

Nous mesurons pour chaque test le temps de calcul nécessaire à son exécution dans des conditions de calcul identiques.

4.6.3.2 Résultats

Le tableau 4.1 présente un résumé des résultats obtenus pour les meilleures paramétrisations de chaque flux de traitement du plan de test (résultats complets en annexe A). À noter qu'il est de loin préférable de sur-segmenter une cellule plutôt que de la sous-segmenter car il est bien plus aisé de procéder à une correction de sur-segmentation.

Les résultats obtenus sur l'image au stade de 162 cellules démontrent que les méthodes basées sur la structure suivies d'une étape de reconstruction par vote de tenseurs sont plus performantes qu'une méthode classique de filtrage morphologique pour le pré-traitement d'images fortement parasitées (internalisation du cytoplasme).

L'analyse des résultats montre également que le pré-traitement proposé dans ce travail aboutit à des résultats significativement améliorés — en particulier dans le cas de la seconde image — et à une vitesse d'exécution largement supérieure aux résultats obtenus via [Mosaliganti 2012] (vote de tenseurs 30 fois plus rapide pour une initialisation des primitives avec une image de plans centraux binarisés).

Ces résultats abondent dans le sens d'une utilisation du procédé de votes de tenseurs sur des données binarisées et éparées. En effet, une initialisation dense de l'image des tenseurs votants — telle que proposée dans [Mosaliganti 2012] — génère beaucoup d'informations autour des plans centraux de membranes. Au delà du facteur de temps de calcul, nous montrons grâce à ces tests que cette information supplémentaire est au mieux inutile, au pire nuisible en raison d'un effet non négligeable de « télescopage » des votes de tenseurs provenant de points à la fois éloignés des plans centraux et mal orientés spatialement.

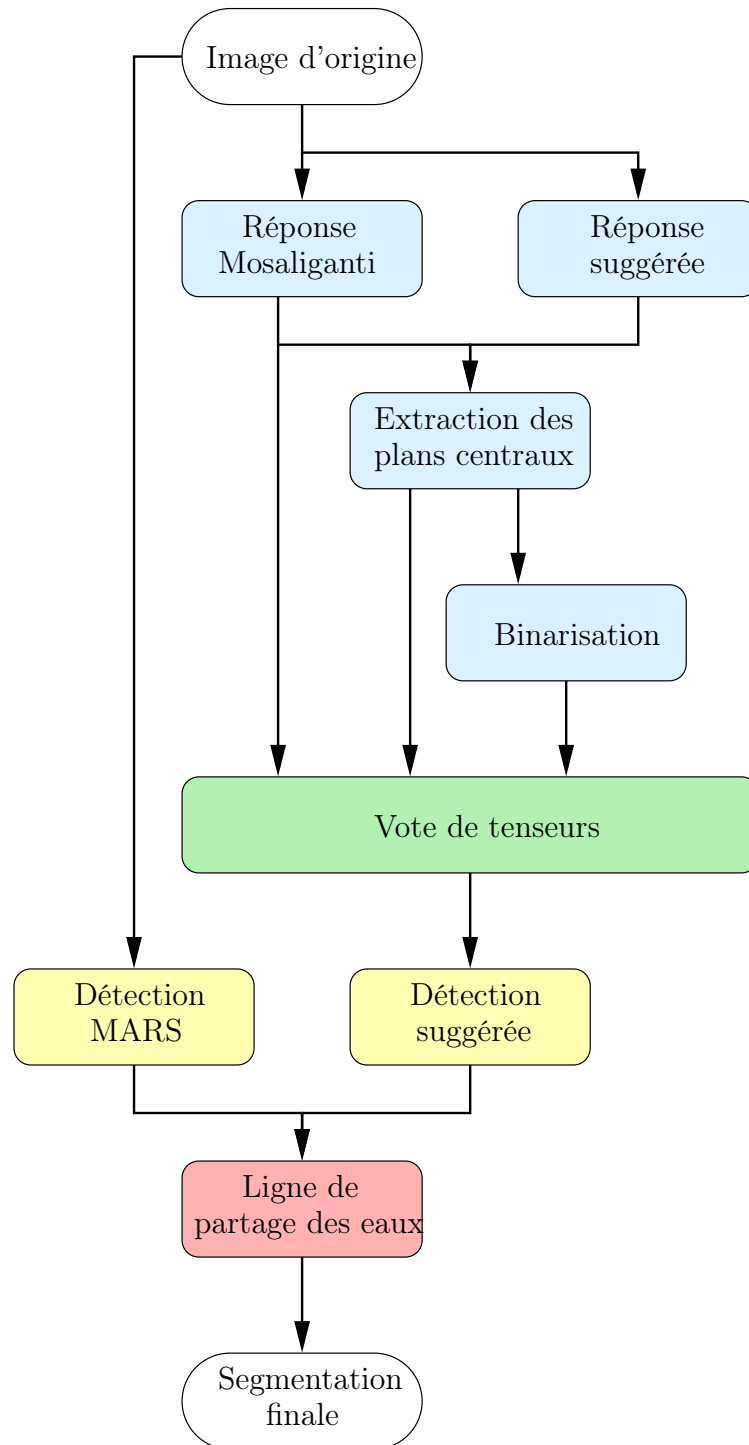


FIGURE 4.16 – Plan de tests mis en place pour l'évaluation des méthodes de détection et de reconstruction des membranes. En bleu : détection des membranes. En vert : reconstruction d'une image en niveaux de gris. En jaune : détection des cellules. En rouge : segmentation des cellules.

Image (#cellules)	Réponse	Vote de tenseurs	σ	h_{\min}	TD	OD	UD	Temps VT (s)
Im1 (32)	Michelin	BINAIRE		30	30	2	0	1943
	Michelin	EXTREMA		10	28	4	0	8253
	Mosaliganti	BINAIRE	5	30	30	2	0	2104
	Mosaliganti	REPONSE	3	10	28	4	0	65210
	MARS	[Fernandez 2010a]			31	9	0	
Im2 (162)	Michelin	BINAIRE		15	149	15	3	2422
	Michelin	EXTREMA		10	144	11	8	7009
	Mosaliganti	BINAIRE	4	30	141	20	4	2965
	Mosaliganti	REPONSE	4	15	139	18	8	64682
	MARS	[Fernandez 2010a]			128	40	5	

TABLE 4.1 – Résultats des détectons pour deux images d’un embryon à différents moments de son développement. #cellules : nombre de cellules de la vérité terrain. Réponse {Michelin/Mosaliganti} : choix de la fonction de planéité. Vote de tenseurs {REPONSE/EXTREMA/BINAIRE} : choix de l’initialisation des primitives (EXTREMA : plans centraux). σ : échelle utilisée (en pixels) pour la fonction de planéité (seulement pour la réponse de Mosaliganti). h_{\min} : paramètre variable de la phase de détection (voir section 4.5). TD/OD/UD : nombre de bonnes/sur-/sous-détectons. Temps VT : temps mis pour effectuer l’étape de vote de tenseurs (en secondes).

Enfin, la fonction de réponse R^n (équation 4.7) que nous proposons suivie d’une extraction et d’une binarisation des plans centraux aboutit à une légère amélioration des résultats de segmentation dans le cas de la seconde image par rapport à la fonction de planéité R_{Mos} (équation 4.2).

4.6.4 Saillance membranaire

La dernière contribution que nous apportons porte sur l’utilisation d’un nouvel extracteur de membranes \mathcal{M} (équation 4.21) favorisant la reconstruction du signal au niveau des jonctions de membranes par rapport à l’extracteur de surfaces \mathcal{S} utilisé dans [Mosaliganti 2012, Martinez-Sanchez 2014]. Nous utilisons les mêmes données de test pour l’évaluation de l’utilisation la saillance membranaire \mathcal{M} à la place de la saillance surfacique \mathcal{S} que pour les précédentes évaluations.

4.6.4.1 Méthodologie

De la même façon que lors de la précédente expérience, nous évaluons quantitativement les segmentations obtenues, pour un pré-traitement identique, à partir des images extraites \mathcal{M} et \mathcal{S} . Nous mesurons le nombre de bonnes, sur- et sous-détectons (TD, OD et UD) de cellules sur la base d’une vérité terrain experte.

Image (#cellules)	Extracteur	TD	OD	UD
Im1 (32)	\mathcal{M}	32	0	0
	\mathcal{S}	30	2	0
Im2 (162)	\mathcal{M}	155	9	2
	\mathcal{S}	149	15	3

TABLE 4.2 – Résultats des détections de cellules : évaluation de l’extracteur de saillance membranaire \mathcal{M} par rapport à l’extracteur de saillance surfacique \mathcal{S} . TD/OD/UD : nombre de bonnes/sur-/sous-détections.

4.6.4.2 Résultats

Le tableau 4.2 présente les résultats de détection des cellules sur les deux images tests selon le choix du type d’extracteur (\mathcal{M} ou \mathcal{S}) après l’étape de vote de tenseurs. On note dans chaque image une amélioration significative de la qualité de la segmentation avec l’extracteur de saillance membranaire que nous proposons. La raison de cette amélioration est qu’un défaut fréquent avec l’utilisation de \mathcal{S} est la détection de graines parasites au niveau de jonctions entre groupes de cellules. Cette détection est due à l’affaiblissement du signal reconstruit en ces régions, faisant apparaître des minima régionaux qui donneront lieu à des régions segmentées après application de la ligne de partage des eaux. La force de l’extracteur \mathcal{M} est de bien mieux appréhender la reconstruction du signal au niveau des jonctions.

De plus, la segmentation des cellules obtenue avec l’utilisation de l’extracteur \mathcal{M} présente une forme qualitativement nettement plus fidèle à celle attendue qu’avec l’utilisation de \mathcal{S} au niveau des jonctions de cellules (figure 4.17).

4.7 Conclusions et perspectives

Dans ce chapitre, nous avons présenté une chaîne de traitement pour la segmentation d’images de tissus épithéliaux marqués au niveau des membranes. Cette chaîne de traitement reprend dans les grandes lignes les mêmes étapes qu’une autre méthode proposée par [Mosaliganti 2012] (figure 4.18) qui a prouvé son efficacité sur des données au faible rapport signal-sur-bruit.

Nous apportons plusieurs contributions originales à cette chaîne de traitement. Nous proposons une nouvelle fonction de planéité basée structure et multi-échelle. Nous procédons à l’extraction et au seuillage des plans centraux issus de la fonction de réponse de manière automatisée avec une méthode de seuillage adaptatif tenant compte de l’anisotropie du signal des membranes induite par le système d’imagerie. Nous utilisons le résultat de la binarisation des plans centraux à la place de la fonction de réponse pour le choix des primitives du procédé de vote de tenseurs, ce qui permet d’éliminer de l’information nuisible à la bonne reconstruction des membranes tout en divisant le temps de calcul du vote de tenseurs par plus de 30. Enfin, nous proposons un nouvel extracteur de « saillance membranaire » permettant une

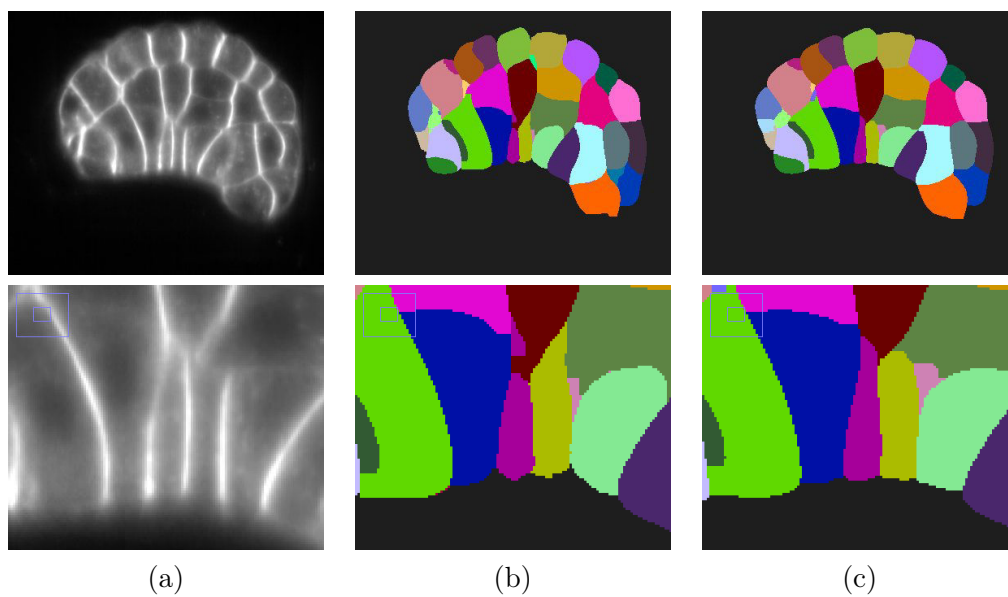


FIGURE 4.17 – Comparaison du résultat de la segmentation d'une image selon le choix de l'extracteur. Ici, le marquage des régions pour la ligne de partage des eaux est identique et nous ne comparons que les différences de morphologie des régions segmentées. (a) Image d'origine (vue en coupe). (b) Image segmentée avec l'extracteur \mathcal{S} . (c) Image segmentée avec l'extracteur \mathcal{M} . Première ligne : vue en coupe entière. Deuxième ligne : zoom sur une région de la coupe.

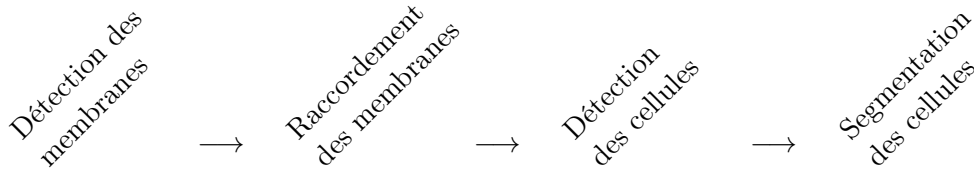


FIGURE 4.18 – Chaîne de traitement pour la segmentation des cellules.

meilleure reconstruction du signal au niveau des jonctions entre plusieurs cellules.

L'ensemble de ces contributions apporte de nouvelles alternatives pour la numérisation d'images d'organismes en développement. Nous avons testé cette chaîne de traitement sur des données de *Phallusia mammillata* et nous avons montré quantitativement que les performances de segmentation sont améliorées par rapport à la méthode d'origine [Mosaliganti 2012] et par rapport à la méthode de référence MARS [Fernandez 2010a] utilisée par [Guignard 2015] sur ce jeu de données.

Il existe de nombreuses perspectives pour ce travail. Il est tout d'abord essentiel d'aborder l'aspect dynamique du développement des tissus, c'est-à-dire de numériser des séquences $3D+t$ incluant le suivi des lignées cellulaires. De plus, des outils d'extraction d'informations quantitatives sur un tissu numérisé (taille, forme des cellules, surfaces d'interaction entre cellules) peuvent être intégrés à la méthode développée. Enfin, des tests sur des données différentes (images de méristèmes d'arabidopsis) ont été réalisés mais soulèvent de nouveaux problèmes liés à la technique d'imagerie : l'étape de binarisation des plans centraux doit être ajustée selon les spécificités du système d'acquisition des images.

Toutes les étapes de la méthode sont implémentées en langage C dans un environnement Unix.

Recalage d'embryons d'ascidies

Sommaire

5.1	Introduction	69
5.2	Recalage d'embryons segmentés : notre approche	70
5.3	Détection de la symétrie bilatérale de l'embryon	74
5.3.1	Extraction des axes de symétrie bilatérale candidats	76
5.3.2	Initialisation des plans de symétrie candidats	76
5.3.3	Sélection du meilleur plan candidat	78
5.3.4	Ajustement de l'interface de symétrie bilatérale par un plan	78
5.4	Recalage d'embryons d'ascidies	79
5.4.1	Initialisation des appariements	81
5.4.2	Itérations : choix de la fonction de coût	82
5.4.3	Solution globale	83
5.5	Expériences et résultats	83
5.5.1	Données	83
5.5.2	Extraction du plan de symétrie bilatérale	85
5.5.3	Recalage d'embryons	88
5.6	Conclusions et perspectives	91

5.1 Introduction

L'un des principaux défis de la biologie développementale est de parvenir à décrire le processus de morphogenèse, c'est-à-dire comment chaque tissu d'un organisme en développement évolue jusqu'à atteindre sa forme fonctionnelle selon l'expression génique de l'organisme. L'imagerie est la méthode de choix pour l'étude de ce processus [Megason 2007]. Les techniques de microscopie récentes permettent aujourd'hui de réaliser des observations d'organismes en développement avec une résolution spatio-temporelle suffisamment fine pour permettre l'observation du développement d'embryons entiers à l'échelle sub-cellulaire [Keller 2013].

En analyse d'image, le développement de méthodes de recalage permettant la comparaison d'individus ou l'agrégation de données provenant de multiples individus constitue un domaine de recherche important. Le recalage d'individus prodigue une solution technique pour plusieurs problèmes, comme la description de la variabilité du développement des organismes en développement au sein d'une même espèce et entre différentes espèces chez l'animal [Munro 2006]. Il existe un besoin de

comparer le développement d'individus observés dans des conditions expérimentales variables et de réaliser des analyses statistiques sur des populations. Une autre motivation pour le développement de méthodes de recalage est la possibilité d'agréger des données provenant de plusieurs individus et de plusieurs modalités d'imagerie dans un même espace afin de corrélérer l'expression génique spatio-temporelle avec les mécanismes de contrôle régulant le comportement des cellules et leurs changements de forme au cours de la morphogenèse [Luengo-Oroz 2011]. Des travaux ont été réalisés pour proposer des stratégies de recalage dédiés à l'embryogenèse des drosophiles [Fowlkes 2008], des poissons-zèbres [Castro 2009], des *C. elegans* [Long 2009] et des ascidies [Guignard 2014].

Dans ce chapitre, nous nous focalisons sur l'étude de l'ascidie *Phallusia mammillata*. Le développement très stéréotypé de son embryon fait apparaître des invariances au cours de son développement. À un même stade de développement, il existe une correspondance cellule à cellule entre les embryons sauvages. La capacité à identifier automatiquement les correspondances entre cellules d'embryons permettrait de quantifier la variabilité développementale inter-individuelle à l'échelle de la cellule chez l'embryon de l'ascidie. Dans la suite de ce chapitre, nous proposons une méthode de recalage spatial d'embryons d'ascidies basée sur l'identification de ces correspondances cellule à cellule. À notre connaissance, l'état de l'art ne propose pas de méthode publiée permettant de réaliser cette opération à l'exception de celle que nous développons ci-après.

Organisation du chapitre. Nous rappelons dans la section 5.2 la notion de recalage entre paires d'images. Puis nous présentons la méthode de recalage d'embryons d'ascidies que nous avons développée. Cette méthode se décompose en deux étapes. La première étape (section 5.3) consiste à extraire la symétrie bilatérale (définissant une « gauche » et une « droite » séparés par un plan de symétrie appelé plan sagittal) caractéristique d'un embryon d'ascidie jusqu'au stade de têtard [Lemaire 2009]. La seconde étape (section 5.4) consiste à identifier les correspondances de cellules et à calculer une transformation permettant de superposer spatialement deux images segmentées provenant d'un embryon test et un embryon de référence au même stade de développement. La section 5.5 présente les expériences réalisées pour la validation de notre méthode de recalage. Ces expériences portent sur des données provenant de quatre embryons marqués au niveau de leurs membranes. Des expérimentations sur le recalage multimodal entre un embryon test marqué au niveau de ses noyaux de cellules et un embryon de référence marqué au niveau de ses membranes cellulaires sont également présentées en annexe B. Nous concluons en section 5.6 sur la méthode proposée et discutons de ses perspectives.

5.2 Recalage d'embryons segmentés : notre approche

Les méthodes de recalage en imagerie biologique s'inspirent assez largement de celles appliquées au domaine de l'imagerie médicale, qui constituent un sujet de

recherche très largement étudié [Sotiras 2013]. Dans la plupart des applications, le recalage implique deux images représentant un même objet (une image source ou « test » et une image de référence ou « cible »). Le problème du recalage entre deux images I et J consiste à calculer une transformation optimale $\hat{\mathbf{T}}$ parmi un ensemble de transformations \mathcal{T} qui minimise une fonction de coût \mathcal{C} :

$$\hat{\mathbf{T}} = \arg \min_{\mathbf{T} \in \mathcal{T}} \mathcal{C}(I, J, \mathbf{T}). \quad (5.1)$$

Pour tout problème de recalage, il faut donc déterminer :

- la famille de transformations \mathcal{T} que l'on veut considérer ;
- la fonction de coût \mathcal{C} à minimiser ;
- la stratégie de calcul de la transformation optimale $\hat{\mathbf{T}}$ minimisant la fonction de coût.

Il existe trois principales classes d'outils de recalage : le recalage *iconique*, le recalage *géométrique* et le recalage *hybride*. Les méthodes iconiques sont directement basées sur l'intensité des images avec un critère de similarité à optimiser sur le domaine entier des images. Les méthodes géométriques sont basées sur la mise en correspondance de primitives (points, lignes, surfaces, volumes, vecteurs, repères, etc.) extraites des images source et de référence. Les méthodes hybrides essaient de tirer profit des avantages prodigués par l'utilisation conjointe d'information iconique et géométrique.

Dans le problème que nous étudions, on suppose connues les segmentations des cellules des embryons que l'on souhaite recalculer. Dans ce cas, il est approprié d'envisager une stratégie de recalage géométrique où l'on considère des ensembles de primitives (points) C et D extraites des images segmentées I et J :

$$\begin{aligned} I &\mapsto C = \{c_i\}_{i \in [1, n_C]}, \\ J &\mapsto D = \{d_i\}_{i \in [1, n_D]}. \end{aligned} \quad (5.2)$$

Dans notre application, les points que nous considérons sont les centres de masse des cellules segmentées. Nous émettons l'hypothèse que la localisation du centre de masse d'une cellule donnée est suffisamment stable d'un embryon à l'autre pour fournir une primitive pertinente en terme d'invariance. La formulation du recalage des images I et J devient alors :

$$\hat{\mathbf{T}} = \arg \min_{\mathbf{T} \in \mathcal{T}} \mathcal{C}(C, D, \mathbf{T}). \quad (5.3)$$

Un algorithme de choix pour réaliser la minimisation de la fonction de coût \mathcal{C} est l'algorithme du plus proche point itéré (*Iterative Closest Point*, ICP), introduit par [Besl 1992]. Ce procédé vise à minimiser itérativement la distance entre les nuages de points C et D . À chaque itération i , on détermine la meilleure transformation $\mathbf{T}^{(i)}$ superposant un ensemble transitoire de paires de points $Map^{(i-1)} \in (C \times D)^{n_{i-1}}$ (l'équation 5.3 étant remplacée par l'équation 5.4), classiquement au sens des moindres carrés. n_i dénote ici le nombre d'appariements de points réalisés lors de l'itération i . Remarquons que certains barycentres de cellules peuvent demeurer non appariés et que le nombre d'appariements de points n_i à l'itération i est par conséquent

variable. À partir de cette transformation, un nouvel ensemble de correspondances $Map^{(i)} = Map(\mathbf{T}^{(i)})$ entre les points les plus proches de C et $\mathbf{T}^{(i)}(D)$ est dressé (équation 5.5). Le procédé prend fin lorsque les appariements de points se stabilisent, donnant alors une transformation optimale $\widehat{\mathbf{T}}$ recalant D sur C et un ensemble de points appariés \widehat{Map} associé (algorithme 2). Les appariements de points permettent alors d'identifier les correspondances cellule à cellule entre les images considérées.

$$\forall i \geq 1, \quad \mathbf{T}^{(i)} = \arg \min_{\mathbf{T} \in \mathcal{T}} \mathcal{C} \left(Map^{(i-1)}, \mathbf{T} \right) \quad (5.4)$$

$$\forall (c, d) \in C \times D, \quad (c, d) \in Map(\mathbf{T}) \Leftrightarrow \begin{cases} c = \arg \min_{\tilde{c} \in C} \text{dist}(\tilde{c}, \mathbf{T}(d)) \\ d = \arg \min_{\tilde{d} \in D} \text{dist}(c, \mathbf{T}(\tilde{d})) \end{cases} \quad (5.5)$$

Algorithme 2 Étapes de l'algorithme du plus proche point itéré.

Init $Map^{(0)}$
 $i \leftarrow 0$
répéter
 $i \leftarrow i + 1$
Calculer $\mathbf{T}^{(i)}$
Calculer $Map^{(i)} = Map(\mathbf{T}^{(i)})$
jusqu'à convergence (i.e. $Map^{(i)} = Map^{(i-1)}$)
 $\widehat{Map} \leftarrow Map^{(i)}$
 $\widehat{\mathbf{T}} \leftarrow \mathbf{T}^{(i)}$

L'algorithme du plus proche point itéré débute par une étape d'initialisation où l'on détermine un ensemble d'appariements initial $Map^{(0)}$. Si cette initialisation n'est pas suffisamment proche de la solution optimale \widehat{Map} , l'algorithme du plus proche point itéré peut converger vers une solution locale différente de celle escomptée. L'initialisation de l'algorithme constitue donc une étape très sensible, bien que nécessaire car il n'est pas envisageable de tester exhaustivement toutes les associations de cellules possibles. Une solution est de guider cette étape par l'introduction d'une transformation initiale $\mathbf{T}^{(0)}$ qui rend compte de connaissances *a priori* sur les objets à recalcer sous forme de contraintes géométriques. On définit ainsi l'appariement initial $Map^{(0)} = Map(\mathbf{T}^{(0)})$.

Une caractéristique morphologique des embryons d'ascidies est l'existence d'une symétrie bilatérale à un stade de développement précoce [Nishide 2012] (figures 5.1 et 5.2). Dans la suite de ce chapitre, nous proposons une méthode d'extraction de l'axe de symétrie bilatérale des embryons (section 5.3) et nous initialisons le procédé de recalage des primitives en nous appuyant sur cette connaissance. Nous recalons finalement les ensembles de points C et D et identifions les correspondances entre cellules grâce à une variante robuste de l'algorithme du plus proche point itéré (section 5.4).

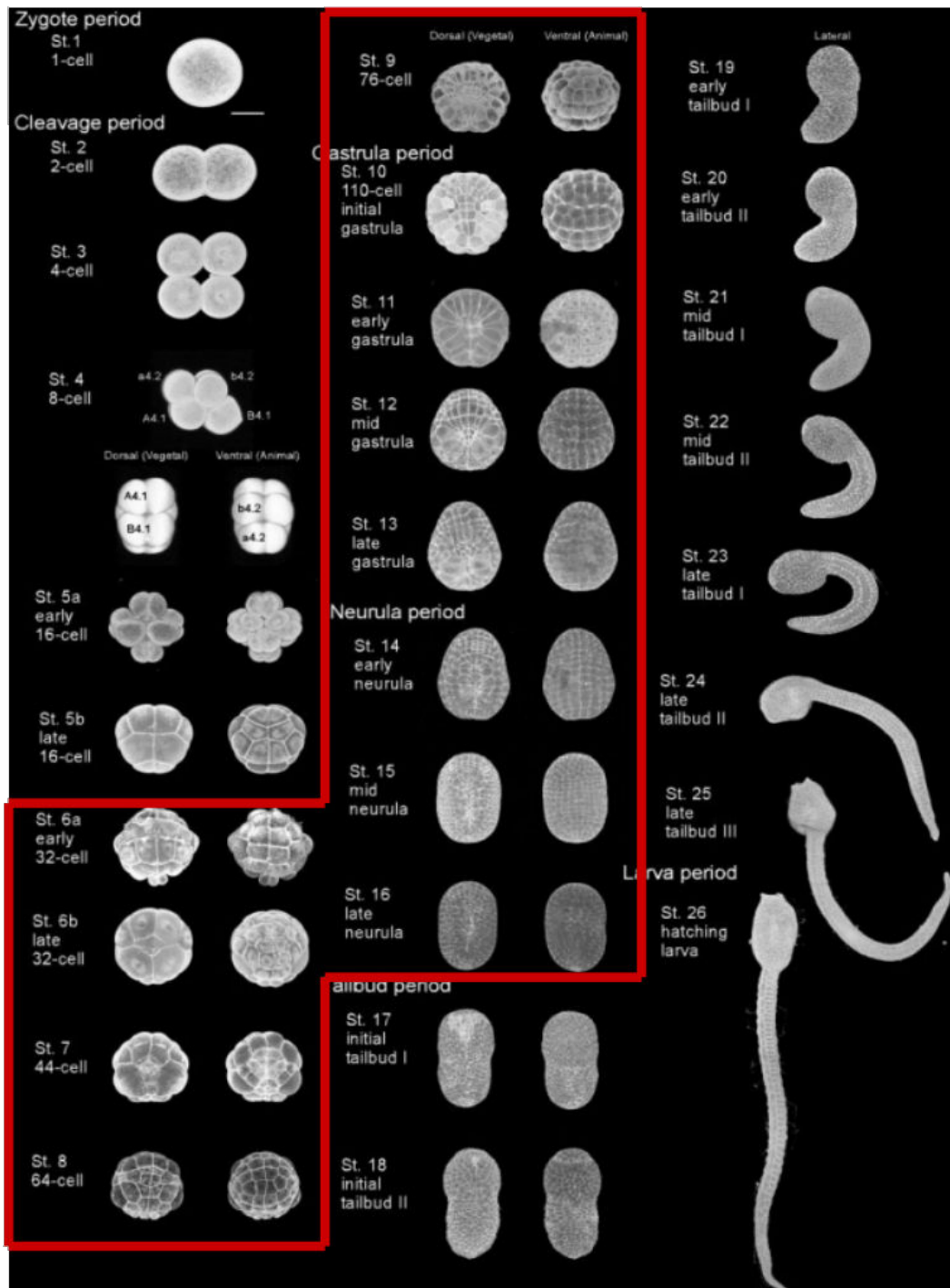


FIGURE 5.1 – Stades de développement de l'ascidie *Ciona intestinalis* du stade d'œuf au stade de larve [Hotta 2007]. Les données que nous traitons proviennent de l'ascidie *Phallusia mammillata*, dont le développement est très similaire. Les stades de développement encadrés en rouge correspondent à ceux des données sur lesquelles nous avons travaillé. La morphologie des embryons fait apparaître une symétrie bilatérale pour ces stades de développement.

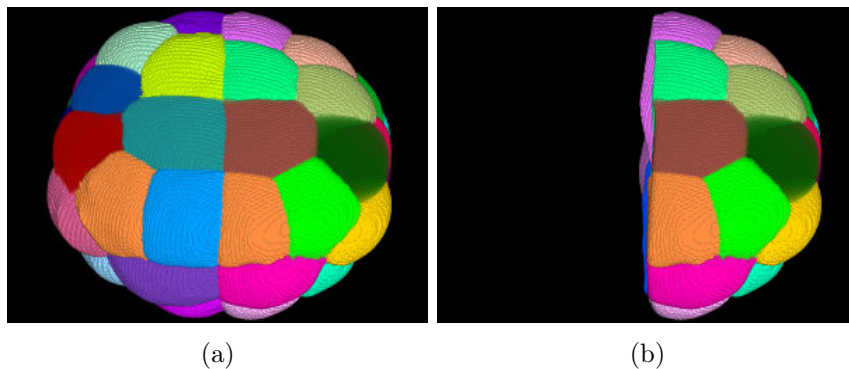


FIGURE 5.2 – Vue 3D d'une image segmentée d'embryon d'ascidie. (a) Embryon complet. (b) Demi-embryon (cellules de la moitié gauche cachées). On voit ici que les membranes séparant les hémisphères gauche et droite de l'embryon constituent une interface assimilable à un plan de symétrie bilatérale de l'organisme.

5.3 Détection de la symétrie bilatérale de l'embryon

Jusqu'au stade neurula, les membranes des cellules à la frontière entre les deux hémisphères de l'embryon, caractérisant la symétrie bilatérale d'un embryon d'ascidie, forment une interface de forme plane (figure 5.2). Nous proposons dans cette section de déterminer l'équation du plan ajustant au mieux cette interface.

Le principe de notre méthode d'extraction du plan de symétrie d'un embryon se décompose en quatre étapes qui s'appuient sur les résultats d'extraction de normales aux membranes et de segmentation calculés au préalable à l'aide de la méthode du chapitre 4 (figure 5.3) :

1. estimation d'un ensemble de directions candidates pouvant porter l'axe de symétrie bilatérale de l'embryon à partir de la distribution des directions des normales aux membranes précédemment calculées ;
2. un plan normal associé à chaque direction candidate est calculé ;
3. le meilleur plan candidat est sélectionné en maximisant un critère de similarité entre les hémisphères gauche et droit de l'embryon binarisé ;
4. l'équation du plan est affinée à l'aide d'un ajustement aux moindres carrés.

Notations. Dans la suite de cette section, on considère une image en niveaux de gris I d'un embryon marqué au niveau de ses membranes cellulaires. B dénote l'image des plans centraux binarisés de la section 4.3, tandis que $\mathbf{v}(P)$ dénote l'orientation normale (c'est-à-dire le vecteur propre associé à la plus large valeur propre de la matrice hessienne en P) associée à un point P binarisé (c'est-à-dire vérifiant $B(P) = 1$). Enfin, S dénote la binarisation de la segmentation finale de I (section 4.5).

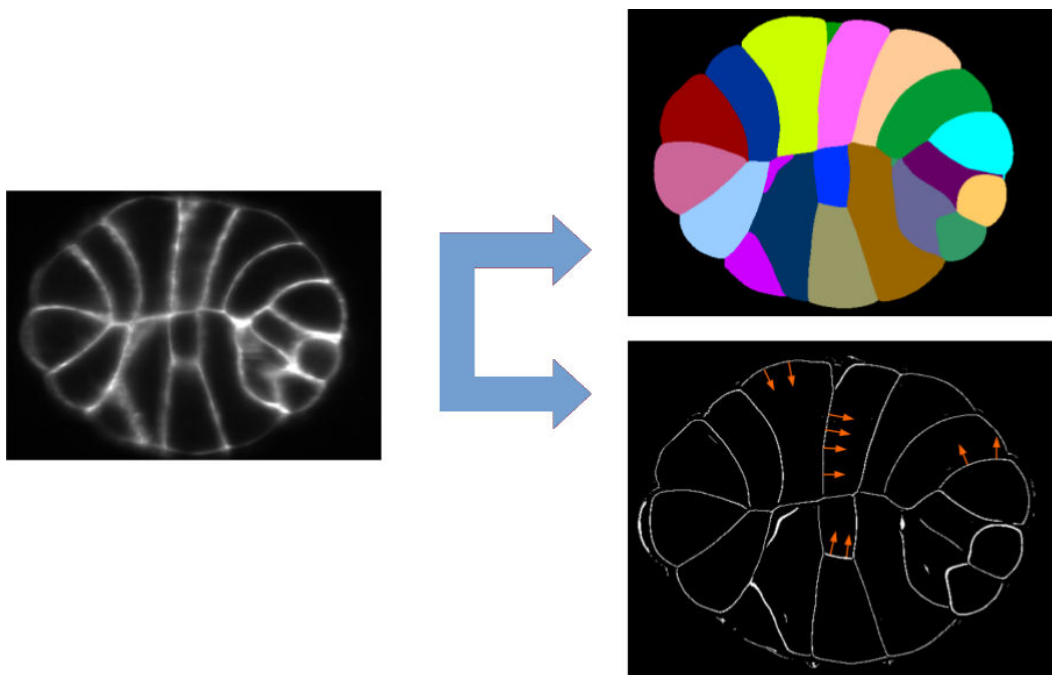


FIGURE 5.3 – Résultats de la segmentation provenant de la chaîne de traitement du chapitre 4. À gauche, vue en coupe d'une image d'intensité d'embryon de *Phallusia mammillata*. En haut à droite, le résultat de la segmentation des cellules de l'image. En bas à droite, le résultat de l'extraction des membranes orientées (section 4.3). Les flèches orange indiquent des directions de normales aux membranes binarisées.

5.3.1 Extraction des axes de symétrie bilatérale candidats

Étant donné la proportion élevée de membranes cellulaires formant la structure plane à l'interface entre les hémisphères gauche et droite de l'embryon, nous formulons l'hypothèse qu'un grand nombre d'orientations normales $\mathbf{v}(P)$ sont alignées avec la normale au plan de symétrie recherché. Ainsi, nous supposons que la distribution des directions normales aux membranes extraites fera apparaître un maximum pour la direction normale au plan de symétrie.

Pour calculer cette distribution, nous discrétisons la sphère unité en un ensemble de vecteurs $\mathbf{N} = \{\mathbf{n}_i\}$ tel que $\forall \mathbf{n}_i \in \mathbf{N}, -\mathbf{n}_i \in \mathbf{N}$. Typiquement, $|\mathbf{N}| \approx 2000$ vecteurs. Nous calculons ensuite une estimation par noyau — également appelée fenêtrage de Parzen [Parzen 1962] — de la densité de probabilité D des directions normales $\mathbf{v}(B)$ sur le support \mathbf{N} :

$$D(\mathbf{n}_i) = \frac{1}{|B|} \times \sum_{P \in B} \frac{1}{N(P)} w(\mathbf{n}_i, \mathbf{v}(P)), \quad (5.6)$$

où $|B|$ est le cardinal de B , $N(P)$ est une constante de normalisation et le noyau $w(\cdot, \cdot)$ est défini par :

$$w(\mathbf{n}, \mathbf{v}) = \exp \frac{-\arccos(|\mathbf{n} \cdot \mathbf{v}|)^2}{2\sigma^2} \quad (5.7)$$

avec $\|\mathbf{n}\| = \|\mathbf{v}\| = 1$.

Le paramètre de fenêtrage σ de l'équation 5.7 détermine le degré de lissage de l'estimation de la densité de probabilité D . Nous fixons $\sigma = 2^\circ$, l'expérience montrant que cette fenêtre permettait une extraction suffisamment fine des axes candidats pour notre problème de détection du plan de symétrie des embryons (figure 5.4).

Soit D_{\max} la valeur maximale globale de D . Nous extrayons l'ensemble $\{\bar{\mathbf{n}}_k\}$ des maxima locaux de D vérifiant $D(\bar{\mathbf{n}}_k) \geq D_{\max}/2$, $\{\bar{\mathbf{n}}_k\} \subset \mathbf{N}$, ce afin de limiter le nombre de directions à tester. Par construction, D est symétrique, et étant donné qu'un plan n'a pas besoin d'orientation, nous ne considérons dans la suite que les maxima de la demi-sphère unité (figure 5.5). $\{\bar{\mathbf{n}}_k\}$ constitue l'ensemble des directions candidates pouvant porter l'axe de symétrie bilatérale de l'embryon.

Note. Afin d'exclure la plupart des fausses détections de membranes du calcul de D , il est possible d'appliquer à B un masque M pouvant être construit à partir des frontières dilatées des cellules segmentées, auquel cas il faut considérer l'image $B \cap M$ en lieu et place de B dans l'équation 5.6 ainsi que dans l'équation 5.11 de la section 5.3.4.

5.3.2 Initialisation des plans de symétrie candidats

Nous définissons l'ensemble des plans de symétrie candidats $\{\mathcal{P}_k\}$ où chaque plan \mathcal{P}_k est paramétré par un vecteur normal $\bar{\mathbf{n}}_k$ précédemment calculé et par un scalaire d_k :

$$\mathcal{P}_k : \bar{\mathbf{n}}_k \cdot [x, y, z] + d_k = 0, \quad (5.8)$$

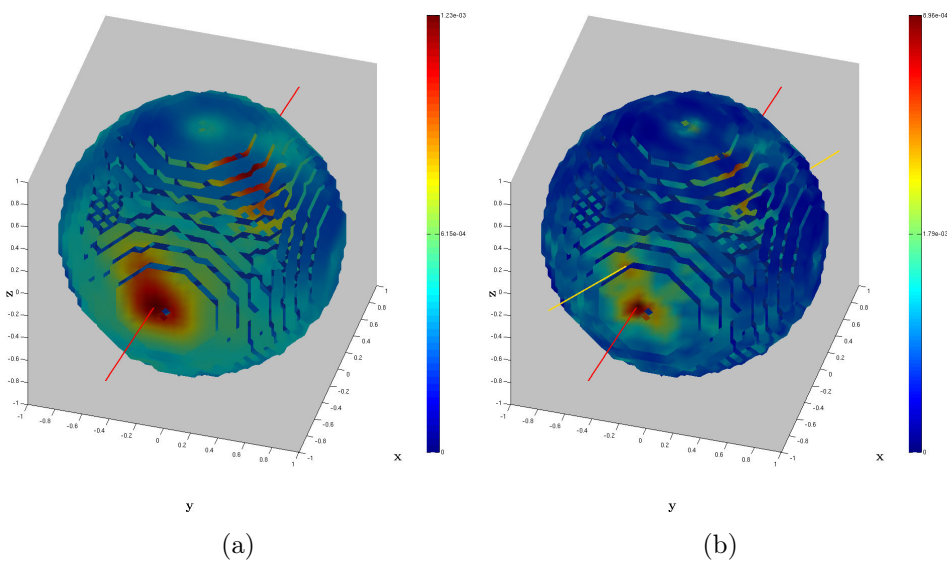


FIGURE 5.4 – Influence du paramètre de fenêtrage sur l'estimation par noyau de la densité de probabilité des directions normales associées aux membranes. (a) $\sigma = 5^\circ$. (b) $\sigma = 2^\circ$. Le choix d'une valeur de σ trop large peut provoquer la fusion de maxima locaux, comme le montre l'extraction des maxima de (a) par rapport à celle de (b).

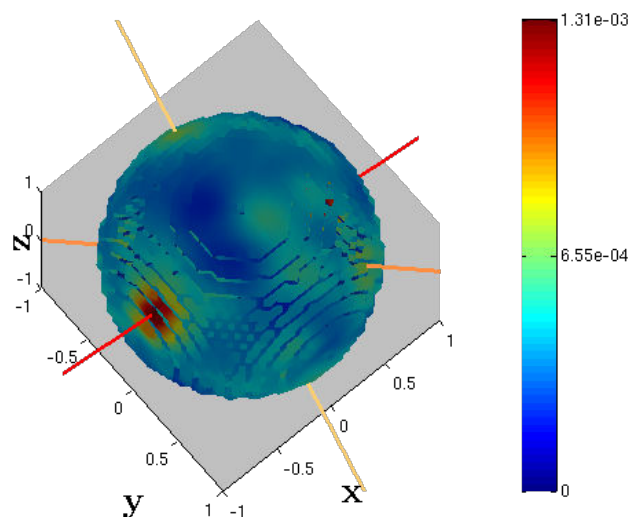


FIGURE 5.5 – Exemple d'estimation par noyau de la densité de probabilité des directions normales associées aux membranes (équation 5.6). L'axe rouge correspond à l'extraction du maximum global de densité. Les autres axes correspondent à des extractions de maxima locaux de densité d'amplitude au moins égale à la moitié de celle du maximum global.

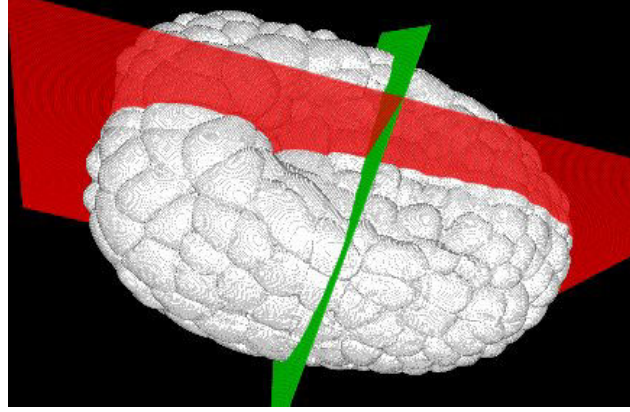


FIGURE 5.6 – Sélection du meilleur plan candidat : exemple sur un embryon au stade de développement neurula avancé. Les mesures de Dice calculés à partir du plan rouge et du plan vert valent respectivement 0,947 et 0,883. Le plan rouge est donc sélectionné.

où d_k est calculé afin que l'embryon segmenté soit équitablement réparti de part et d'autre du plan. Pour cela, nous maximisons l'indice de Dice [Dice 1945] utilisé en tant que mesure de similarité entre la segmentation binarisée S de l'embryon et son symétrique $S \circ \mathcal{P}_{\bar{\mathbf{n}}_k}(d)$ par rapport à un plan $\mathcal{P}_{\bar{\mathbf{n}}_k}(d) : \bar{\mathbf{n}}_k \cdot [x, y, z] + d = 0$ normal à $\bar{\mathbf{n}}_k$, c'est-à-dire pour chaque k :

$$d_k = \arg \max_{d \in \mathbb{R}} \text{Dice}(S, S \circ \mathcal{P}_{\bar{\mathbf{n}}_k}(d)). \quad (5.9)$$

Cette maximisation est menée en pratique sur un ensemble discret borné $\mathcal{D} \subset \mathbb{R}$ de valeurs du paramètre d . La valeur de d_k est ainsi estimée avec une précision typiquement de l'ordre de la taille d'un voxel (i.e. environ $1\mu\text{m}$). Cette précision est en pratique suffisamment fine pour ne pas avoir d'influence sur la sélection du meilleur plan candidat (section 5.3.3) ni sur le résultat de l'ajustement de l'équation du plan de symétrie bilatérale présenté en section 5.3.4.

5.3.3 Sélection du meilleur plan candidat

La sélection du meilleur plan candidat, illustrée en figure 5.6, repose sur la maximisation de la même mesure de similarité que pour l'étape d'initialisation des plans de symétrie candidats (section 5.3.2). Soit \hat{k} l'indice du meilleur plan candidat $\mathcal{P}_{\hat{k}}$:

$$\hat{k} = \arg \max_k \text{Dice}(S, S \circ \mathcal{P}_k). \quad (5.10)$$

5.3.4 Ajustement de l'interface de symétrie bilatérale par un plan

Les expériences montrent que le plan $\mathcal{P}_{\hat{k}}$ calculé en section 5.3.3 est une approximation plutôt bonne de la symétrie bilatérale de l'embryon mais qu'il n'est pas

forcément bien aligné avec les membranes constituant l'interface entre les hémisphères gauche et droite de l'embryon. Le calcul de la normale $\bar{\mathbf{n}}_{\hat{k}}$ à $\mathcal{P}_{\hat{k}}$ est en effet soumis à plusieurs biais. La discrétisation du support \mathbf{N} utilisé pour le calcul de la densité de probabilité D des directions normales aux membranes segmentées introduit une perte de précision pour l'extraction des maxima locaux. Surtout, le calcul de D prend en compte tous les points P tels que $B(P) = 1$ (équation 5.6), ceux n'appartenant pas à l'interface de symétrie pouvant donc perturber davantage la précision de la direction extraite $\bar{\mathbf{n}}_{\hat{k}}$.

Pour une estimation plus précise de l'équation du plan de symétrie bilatérale de l'embryon $\hat{\mathcal{P}}$, nous réalisons un raffinement itératif aux moindres carrés par un plan des points de B situés à l'interface entre les deux hémisphères de l'embryon. La pondération des points de B , mise à jour entre chaque itération, pénalise les points éloignés du plan calculé à l'itération précédente et ceux dont la normale forme un angle important avec celle de ce plan (figure 5.7). Nous définissons donc une série de plans $(\mathcal{P}^{(i)})_{i \geq 0} \xrightarrow{i \rightarrow \infty} \hat{\mathcal{P}}$ initialisée par $\mathcal{P}^{(0)} = \mathcal{P}_{\hat{k}}$ et définie par :

$$\mathcal{P}^{(i+1)} = \arg \min_{\mathcal{P}} \sum_{P \in B} \rho_{\mathcal{P}^{(i)}}(P) \|\mathcal{P}(P)\|^2 \quad (5.11)$$

où la pondération ρ relative à un plan \mathcal{P} est définie par :

$$\rho_{\mathcal{P}}(P) = \exp \frac{-\mathcal{P}(P)^2}{2\sigma_d^2} \exp \frac{-\arccos(|\mathbf{n} \cdot \mathbf{v}(P)|)^2}{2\sigma_a^2}. \quad (5.12)$$

Les paramètres σ_d et σ_a régissent respectivement la pondération de la distance entre le point P de B et le plan \mathcal{P} et la pondération de la différence d'angle entre $\mathbf{v}(P)$ et la normale au plan \mathcal{P} . Nous fixons la valeur de σ_d à la moitié du diamètre approximatif d'une cellule d'embryon et celle de σ_a à 2° . La figure 5.8 présente l'initialisation et le résultat des deux premières itérations de l'ajustement du plan de symétrie bilatérale d'un embryon. Un exemple de résultat de raffinement du plan de symétrie est montré en figure 5.9.

5.4 Recalage d'embryons d'ascidies

Jusqu'au stade de gastrulation, les embryons d'ascidies ont un développement très stéréotypé et invariant. Il est ainsi possible de dresser des correspondances cellule à cellule entre deux embryons à un même stade de développement. L'identification de ces correspondances offre un puissant moyen pour réaliser des études de populations à l'échelle de la cellule, mais est difficile à établir. En effet, à des stades de développement précoces, un embryon ressemble à une sphère carrelée par des cellules, et l'appariement de cellules doit être réalisée par des chercheurs expérimentés. Même après la gastrulation, lorsqu'une concavité (le blastopore) apparaît et permet d'identifier aisément l'orientation d'un embryon, les mises en correspondance de cellules sont rendues complexes en raison du nombre croissant de cellules. Nous proposons dans cette section une méthode de recalage géométrique qui détermine

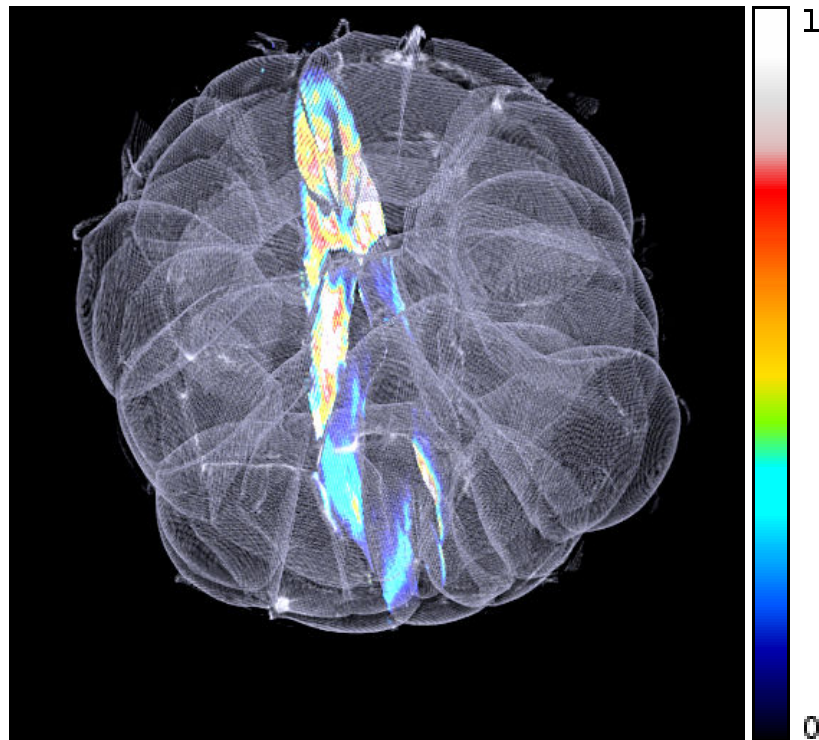


FIGURE 5.7 – Vue 3D en semi-transparence de l'image B des membranes binarisées d'un embryon et en couleur de l'attribution des coefficients de pondération des points de B par l'équation 5.12. Les couleurs vives correspondent à de fortes pondérations, les couleurs sombres à de faibles pondérations, l'absence de couleur à une pondération quasi-nulle.

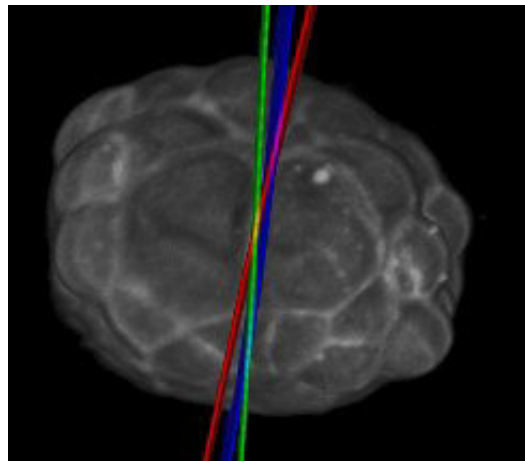


FIGURE 5.8 – Résultat des premières itérations de l'ajustement de l'interface de symétrie bilatérale d'un embryon. Le plan rouge (respectivement bleu, vert) constitue le plan initial (respectivement le premier, le deuxième ajustement).

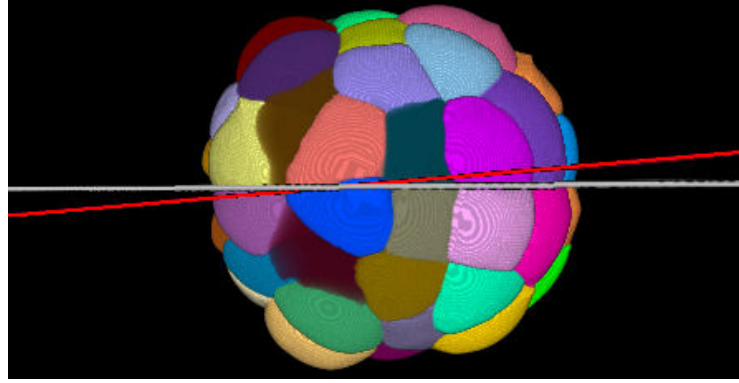


FIGURE 5.9 – Initialisation du plan de symétrie (en rouge) et résultat de son ajustement (en blanc) sur un embryon segmenté au stade de développement 64 cellules.

une transformation affine superposant un embryon test sur un embryon de référence et qui identifie un ensemble de correspondances cellule à cellule. Les segmentations des cellules ainsi que le plan de symétrie bilatérale sont supposés connus pour chaque embryon.

Notations. Dans la suite de la section, I et J désignent respectivement un embryon de référence et un embryon test à un même stade de développement, et nous adoptons les notations suivantes :

- \mathcal{P} et \mathcal{Q} désignent les plans de symétrie bilatérale respectifs de I et J ;
- $\mathbf{n}_\mathcal{P}$ et $\mathbf{n}_\mathcal{Q}$ désignent les vecteurs unitaires normaux à \mathcal{P} et \mathcal{Q} ;
- $C = \{c_i\}_{i \in [1, n_C]}$ et $D = \{d_i\}_{i \in [1, n_D]}$ sont les centres de masse des cellules segmentées de I et J ;
- G et H sont les centres de masse des segmentations entières des embryons I et J ;
- g et h sont les projections respectives de G et H sur les plans \mathcal{P} et \mathcal{Q} .

Pour recalcer les embryons I et J , nous cherchons à calculer la transformation $\hat{\mathbf{T}}$ optimisant un ensemble d'appariements de barycentres de cellules \widehat{Map} au sens d'une fonction de coût \mathcal{C} à l'aide d'une méthode de recalage itérative, l'algorithme du plus proche point itéré (décrit en section 5.2). Dans la suite, \mathcal{T} décrit l'ensemble des transformations affines dans lequel nous recherchons la transformation optimale $\hat{\mathbf{T}}$.

5.4.1 Initialisation des appariements

Pour l'initialisation du problème de recalage, il est déraisonnable d'un point de vue calculatoire de tester tous les ensembles d'appariements possibles entre les points de C et D . L'introduction de contraintes géométriques doit permettre de réduire fortement l'ensemble des appariements initiaux à tester en construisant un ensemble de transformations initiales \mathcal{T}_0 recalant approximativement D sur C afin

d'estimer les ensembles d'appariements initiaux $Map(\mathcal{T}_0) = \{Map(\mathbf{T})\}_{\mathbf{T} \in \mathcal{T}_0}$ selon le critère du plus proche voisin défini par l'équation 5.5.

Nous contraignons l'initialisation du procédé de recalage en imposant le choix de transformations rigides superposant les plans de symétrie bilatérale \mathcal{P} et \mathcal{Q} des embryons, ainsi que la superposition des projections g et h des centres de masse des segmentations sur ces plans de symétrie. Ces contraintes permettent d'obtenir un alignement raisonnable de I et J soumis à une incertitude portant sur l'orientation de \mathbf{n}_Q par rapport à \mathbf{n}_P et à un degré de liberté correspondant à une rotation de l'embryon test autour de son axe de symétrie et centrée en h :

$$\forall \mathbf{T} \in \mathcal{T}, \quad \mathbf{T} \in \mathcal{T}_0 \Leftrightarrow \begin{cases} \mathbf{T} \text{ est une transformation rigide,} \\ \mathbf{T}(h) = g, \\ \mathbf{T}(\mathbf{n}_Q) = \pm \mathbf{n}_P. \end{cases} \quad (5.13)$$

Dans la suite, \mathcal{T}_0^+ (respectivement \mathcal{T}_0^-) désigne le sous-ensemble des transformations $\mathbf{T} \in \mathcal{T}_0$ telles que $\mathbf{T}(\mathbf{n}_Q) = \mathbf{n}_P$ (respectivement $\mathbf{T}(\mathbf{n}_Q) = -\mathbf{n}_P$). En pratique, nous discrétisons \mathcal{T}_0^+ et \mathcal{T}_0^- en n éléments selon des angles de rotation équirépartis entre 0 et 2π radians (typiquement, $n = 128$). On indexe $\mathcal{T}_0^+ = \{\mathbf{T}_k\}_{k \in [1, n]}$ et $\mathcal{T}_0^- = \{\mathbf{T}_{-k}\}_{k \in [1, n]}$.

Pour chaque élément $\mathbf{T}_k \in \mathcal{T}_0^+$ (respectivement $\mathbf{T}_{-k} \in \mathcal{T}_0^-$), nous calculons les ensembles d'appariements de points $Map_k^{(0)} = Map(\mathbf{T}_k)$ (respectivement $Map_{-k}^{(0)} = Map(\mathbf{T}_{-k})$). Dans la suite, chacun de ces ensemble d'appariements de points de C et D sera utilisé comme une initialisation pour l'algorithme du plus proche point itéré (algorithme 2).

5.4.2 Itérations : choix de la fonction de coût

Nous considérons une transformation initiale $\mathbf{T}_k^{(0)}$ ainsi que l'ensemble d'appariements de points correspondant $Map_k^{(0)}$ définis en section 5.4.1, où $|k| \in [1, n]$. Nous définissons la séquence de transformations affines $\{\mathbf{T}_k^{(i)}\}_{i \geq 1} \in \mathcal{T}^{\mathbb{N}^*}$ ainsi que la séquence d'ensembles d'appariements correspondants $\{Map_k^{(i)}\}_{i \geq 1}$ à l'aide de l'algorithme 2 du plus proche point itéré, où l'on calcule une transformation optimale à chaque itération au sens des moindres carrés tamisés [Rousseeuw 1987] :

$$\mathbf{T}_k^{(i)} = \arg \min_{\mathbf{T} \in \mathcal{T}} \mathcal{C}_\alpha \left(Map_k^{(i-1)}, \mathbf{T} \right), \quad (5.14)$$

où la fonction de coût \mathcal{C}_α calcule la moyenne des carrés des distances résiduelles entre points appariés sur un sous-ensemble $\alpha(Map)$ extrait dynamiquement d'un ensemble d'appariements Map excluant les appariements les plus aberrants (au sens d'un large résidu) selon une proportion fixée p_α d'appariements conservés :

$$\mathcal{C}_\alpha (Map, \mathbf{T}) = \frac{1}{\#\alpha(Map)} \sum_{(c, d) \in \alpha(Map)} \text{dist}(c, \mathbf{T}(d))^2. \quad (5.15)$$

Ici, $\#$ désigne l'opérateur cardinal. Le choix de la proportion d'appariements conservés $p_\alpha \in]0, 1]$ (1 correspondant au cas d'une distance résiduelle moyenne non

tamisée) permet d'exclure du calcul de chaque transformation $\mathbf{T}_k^{(i)}$ un certain nombre d'appariements erronés pouvant provenir d'erreurs de segmentation dans I et J . Ce choix doit donc être fait en fonction du nombre estimé d'erreurs de segmentations composant C et D . Nous choisissons dans ce travail de conserver 80% des appariements pour le calcul de la fonction de coût, c'est-à-dire de fixer p_α à 0,8. Ce choix est motivé par le fait que dans nos expériences, les erreurs de segmentation des images que nous considérons ne dépassent pas un taux de 10% par rapport au nombre total de cellules des embryons I et J .

Après convergence de l'algorithme du plus proche point itéré, nous notons respectivement $\widehat{\mathbf{T}}_k$ et \widehat{Map}_k la transformation optimale recalant D sur C et l'ensemble de points appariés associé que l'on obtient en initialisant l'algorithme avec l'ensemble de paires de points $Map_k^{(0)}$.

5.4.3 Solution globale

La transformation optimale $\widehat{\mathbf{T}}$ et l'ensemble d'appariements cellule à cellule optimal \widehat{Map} constituant la solution globale au problème d'optimisation posé en section 5.2 sont ceux minimisant le critère de résidu moyen tamisé (équation 5.15) parmi l'ensemble de solutions optimales obtenues selon les différentes initialisations du problème (figure 5.10) :

$$\begin{cases} \widehat{\mathbf{T}} = \widehat{\mathbf{T}}_{\hat{k}} \\ \widehat{Map} = \widehat{Map}_{\hat{k}} \end{cases} \quad \text{avec } \hat{k} = \arg \min_{|k| \in [1, n]} \mathcal{C}_\alpha(\widehat{Map}_k, \widehat{\mathbf{T}}_k). \quad (5.16)$$

La figure 5.11 présente un exemple de résultat final d'appariements de cellules entre deux embryons au stade de développement 64 cellules.

5.5 Expériences et résultats

5.5.1 Données

Nous avons travaillé sur des séquences temporelles d'images 3D provenant de 5 embryons de *Phallusia mammillata* pour un total de 115 images acquises en utilisant un microscope MuViSPIM (microscope à feuille de lumière multi-angle) [Krzic 2012] :

- 1^{er} embryon : 99 images du stade 32 cellules au stade 172 cellules (gastrulation) ;
- 2^e embryon : 9 images aux stades 64 et 112 cellules ;
- 3^e embryon : 5 images du stade 64 cellules au stade 70 cellules ;
- 4^e embryon : 1 image au stade 114 cellules ;
- 5^e embryon : 1 image au stade neurula (environ 450 cellules).

Les membranes cellulaires des embryons sont marquées par un colorant lipophile (FM4-64) pour le premier embryon et par un marqueur génétique (PH-GFP) pour les autres. Le pas de temps entre deux acquisitions d'images est d'environ une minute. Pour chaque acquisition, quatre images de taille $1200 \times 1200 \times 200$ et de résolution

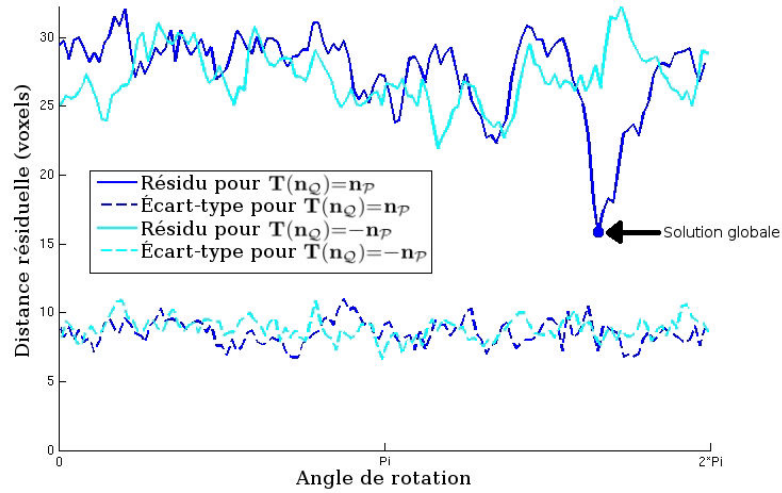


FIGURE 5.10 – Variation de la distance résiduelle moyenne tamisée entre barycentres de cellules appariées après optimisation en fonction de l’initialisation du problème de recalage. La courbe bleue pleine (respectivement cyan pleine) correspond à la variation de la distance résiduelle moyenne selon le paramètre de rotation régissant l’ensemble des transformées initiales \mathcal{T}_0^+ (respectivement \mathcal{T}_0^-). La courbe bleue discontinue (respectivement cyan discontinue) indique la variation de l’écart-type de distance résiduelle selon ce même paramètre de rotation régissant \mathcal{T}_0^+ (respectivement \mathcal{T}_0^-).

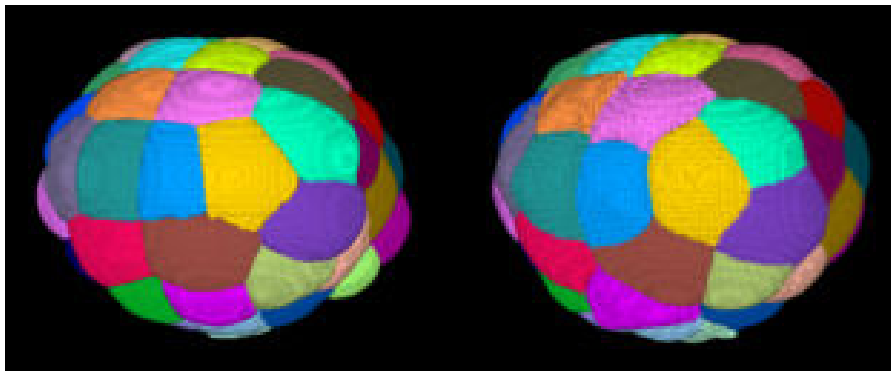


FIGURE 5.11 – Ensemble d’appariements de cellules obtenus entre deux embryons segmentés de *Phallusia mammillata* au stade de développement 64 cellules. Les cellules mises en correspondance apparaissent avec la même couleur.

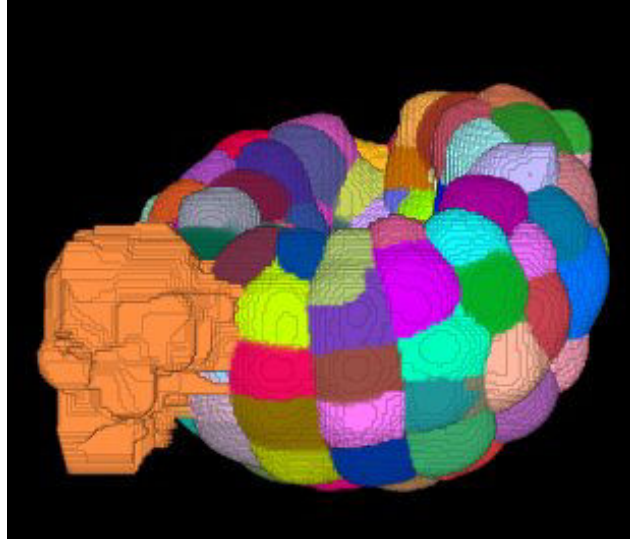


FIGURE 5.12 – Exemple de résultat de segmentation d’une image du premier embryon test. Dans cette segmentation, on remarque l’extraction d’une région aberrante non corrigée (région la plus à gauche en orange).

$0,26 \times 0,26 \times 1\mu\text{m}^3$ provenant de quatre angles de vue séparés de 90° sont construites puis fusionnées. Ainsi, l’image obtenue après fusion possède une résolution isotrope et les défauts d’acquisition relatifs à la microscopie par feuille de lumière sont atténués.

Nous avons appliqué à toutes les images le procédé de segmentation des cellules présenté au chapitre 4. Il est à noter qu’aucune correction *a posteriori* des segmentations n’a été réalisée, ce qui permet de tester la robustesse des méthodes décrites d’extraction du plan de symétrie et de recalage d’individus dans des situations non idéales (figure 5.12).

5.5.2 Extraction du plan de symétrie bilatérale

Nous avons testé la méthode d’extraction du plan de symétrie des embryons d’ascidies présentée en section 5.3 sur les 115 images décrites en section 5.5.1. Nous procédons par un contrôle visuel pour l’évaluation du résultat de l’extraction du plan de symétrie sur une image d’embryon. Bien que purement qualitative, une évaluation experte de l’extraction de la symétrie bilatérale d’un embryon est cependant aisée et rapide à réaliser.

L’inspection visuelle des résultats montre que le plan de symétrie bilatérale a été correctement extrait sur la totalité des images testées. La figure 5.13 présente des exemples d’initialisation et de résultat d’affinage du plan de symétrie sur plusieurs embryons à différents stades de développement, où l’on constate l’apport de l’étape d’ajustement du plan présentée en section 5.3.4. La figure 5.14 présente le résultat de l’extraction du plan de symétrie bilatérale sur chacune des 99 images constituant la séquence temporelle du premier embryon. On observe une consistance convaincante

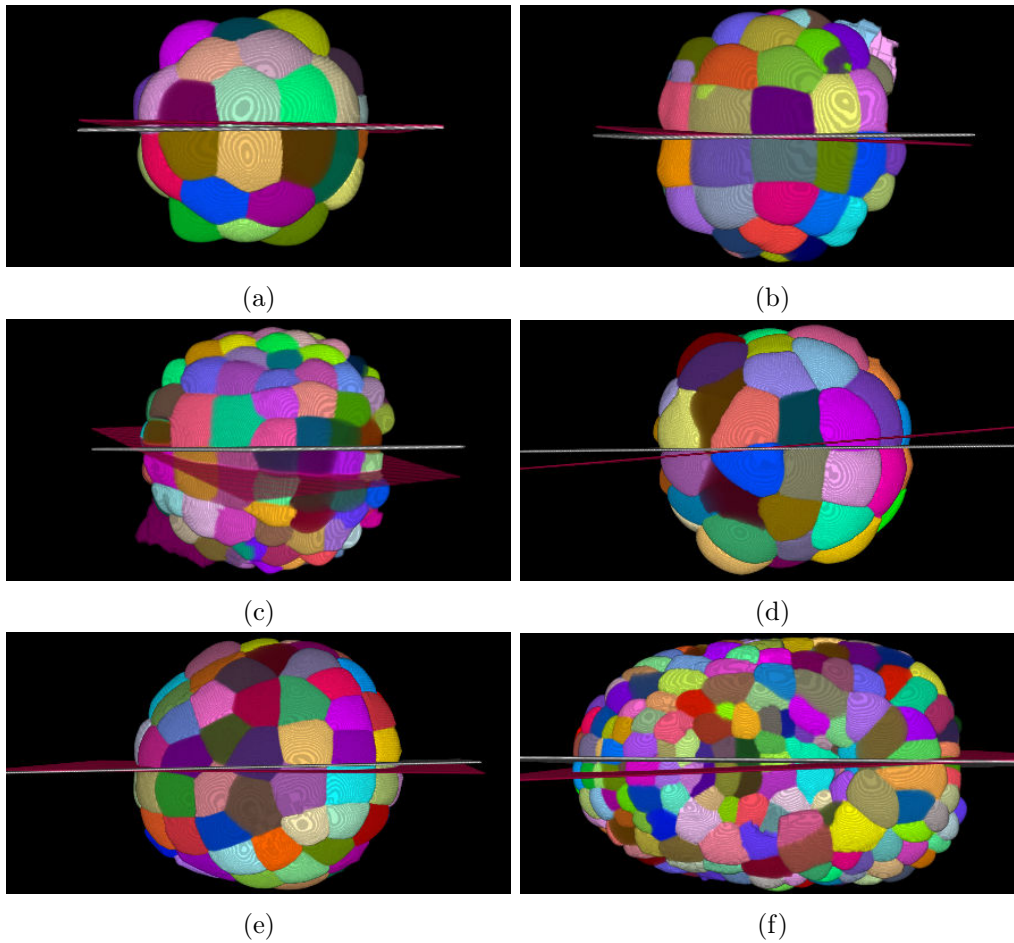


FIGURE 5.13 – Visualisation sur des images segmentées de l'initialisation du plan de symétrie (en rouge) et du plan de symétrie affiné (en blanc). (a-c) 1^{er} embryon aux stades 32 cellules, 76 cellules et gastrula. (d-e) 2^{ème} embryon aux stades 64 et 112 cellules. (f) 5^{ème} embryon au stade neurula.

de la position du plan de symétrie que l'on extrait au cours du temps.

Lors de l'étape d'initialisation, le calcul de la densité de probabilité D des directions normales associées aux membranes sert à estimer des directions spatiales potentiellement porteuses de l'axe de symétrie bilatérale de l'embryon. Les directions extraites correspondent à des maxima locaux de D , notés $\{\bar{\mathbf{n}}_k\}$, que nous trions par valeur décroissante (i.e. $D(\bar{\mathbf{n}}_1) \geq D(\bar{\mathbf{n}}_2) \geq \dots$). L'indice k (appelé ordre) d'un maximum indique donc que ce maximum est le k^e plus grand maximum local de D . L'étape de sélection du meilleur candidat (section 5.3.3) consiste à déterminer l'ordre \hat{k} du maximum de D donnant le plan initial dont on affinera l'équation. Le tableau 5.1 présente les proportions d'ordres de maxima locaux étant sélectionnés pour l'initialisation du procédé d'affinage du plan de symétrie sur les 115 images testées. On observe que la large majorité des maxima sélectionnés sont d'ordre 1, c'est-à-dire correspondent aux maxima globaux de D . Dans les autres cas (soit moins

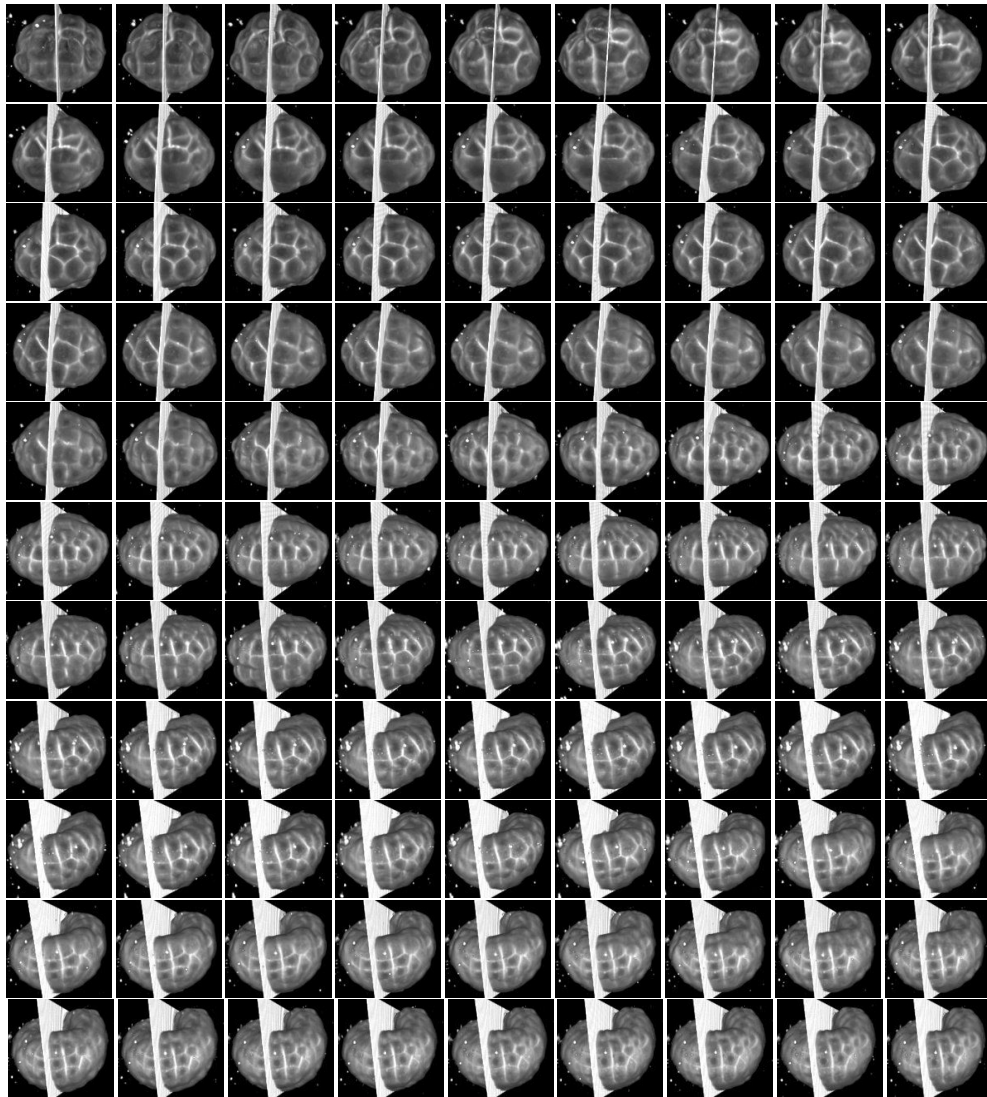


FIGURE 5.14 – Résultat de l'extraction du plan de symétrie bilatérale sur les 99 images originales de la séquence provenant du premier embryon test (vues 3D).

Ordre \hat{k} du maximum local optimal	Proportion sur l'échantillon test
1	86%
2	13%
3	1%

TABLE 5.1 – Proportion de maxima locaux optimaux d'ordre \hat{k} sur l'échantillon test composé de 115 images d'embryons de *Phallusia mammillata*.

de 15% des images de test), l'ordre du maximum est 2 ou 3 au plus. L'investigation de ces cas montre que les maxima locaux de forte intensité formant les plans de symétrie candidats définissent des directions voisines, séparées par des angles inférieurs à 10° (figure 5.4b). Il existe donc une variabilité plus ou moins importante de l'orientation spatiale des membranes autour de l'axe de symétrie bilatérale. Cette étude permet de valider l'hypothèse formulée en début de section 5.3.1 selon laquelle la proportion *a priori* élevée de membranes constituant l'interface de symétrie bilatérale d'un embryon se traduit par l'existence d'un maximum de densité de probabilité des directions normales associées aux membranes selon son axe de symétrie.

La figure 5.15 indique l'histogramme des mesures de similarité entre les embryons binarisés et leur réflexion par rapport au plan de symétrie extrait. Les mesures de Dice sont généralement élevées, avec un indice moyen de 0,9452 et un écart-type de 0,0232. Les plus faibles valeurs de similarité calculées (inférieures à 0,9) sont la conséquence d'erreurs de segmentation qui perturbent la mesure de similarité (voir la figure 5.12). Cependant, le contrôle de ces cas particuliers montre que l'extraction du plan de symétrie est correctement réalisée. Nous montrons ainsi l'importance de considérer une étape d'ajustement local de l'interface de symétrie bilatérale des embryons. L'indice de similarité que l'on maximise en section 5.3.3 considère au contraire une information globale qui permet uniquement de déterminer une position approximative du plan de symétrie, utile à l'initialisation du procédé d'ajustement local.

5.5.3 Recalage d'embryons

La contribution principale de ce chapitre est la proposition d'une méthode automatique de recalage spatial d'embryons d'ascidies permettant d'identifier des correspondances de cellule à cellule chez des individus à un même stade de développement. Notre stratégie d'évaluation consiste à utiliser une image d'embryon de référence à différents stades de développement, à appliquer notre méthode de recalage sur plusieurs images d'embryons tests et à évaluer la qualité et la robustesse de l'identification des appariements de cellules.

- L'embryon de référence (Ref) est le 2^e embryon, avec
 - 1 image au stade 64 cellules (Ref. 64),
 - 1 image au stade 112 cellules (Ref. 112).
- L'embryon test 1 (T1) est le 1^{er} embryon, avec
 - 36 images du stade 36 cellules au stade 76 cellules (T1. 64) recalées sur

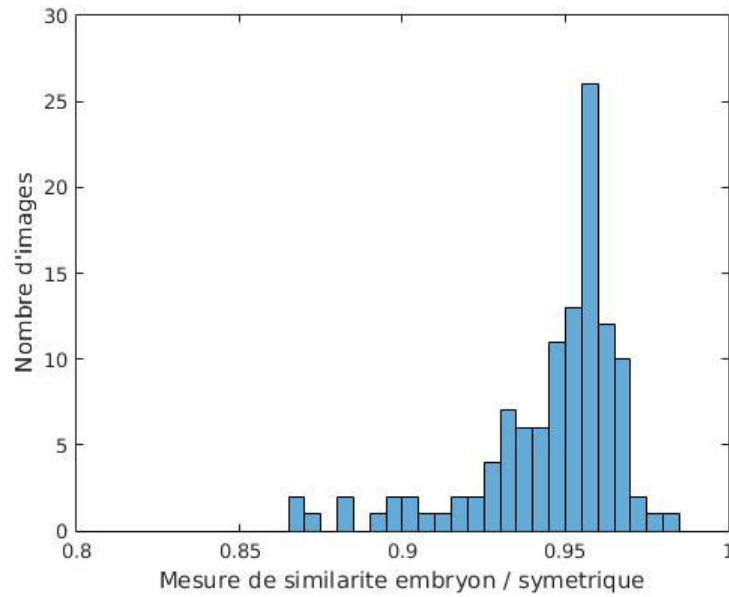


FIGURE 5.15 – Histogramme des mesures de similarité (indice de Dice) entre les embryons segmentés S et leur symétrique $S \circ \hat{\mathcal{P}}$ par rapport au plan de symétrie extrait après ajustement. Mesures portant sur l'ensemble des 115 images de test.

(Ref. 64),

- 45 images du stade 76 cellules au stade 162 cellules (T1. 112) recalées sur (Ref. 112).
- L'embryon test 2 (T2) est le 3^{ème} embryon, avec
 - 5 images du stade 64 au stade 66 cellules (T2. 64) recalées sur (Ref. 64).
- L'embryon test 3 (T3) est le 4^{ème} embryon, avec
 - 1 image au stade 114 cellules recalée sur (Ref. 112).

Notons que dans nos évaluations, nous ne procédons à aucune correction de segmentation sur les embryons que nous recalons afin de réaliser un recalage entièrement automatisé et d'évaluer la robustesse de la méthode proposée dans ces conditions.

Les appariements de cellules sont évalués visuellement. L'inspection visuelle permet de déterminer si le résultat du recalage possède un sens biologique ou non. Pour évaluer la robustesse du calcul de la solution globale entre deux images (section 5.4.3), nous affichons la variation de la distance résiduelle tamisée entre les barycentres des cellules appariées en fonction de l'initialisation du recalage (voir la figure 5.10) pour chacun des recalages effectués. Finalement, nous évaluons la sensibilité de la méthode de recalage aux différences de stades de développement pouvant exister entre l'embryon test et l'embryon de référence. Pour cela, nous recalons chacune des images de la séquence test (T1. 64) sur (Ref. 64) et de la séquence (T1. 112) sur (Ref. 112) — ces séquences englobant respectivement les stades de développement 64 cellules et 112 cellules de l'embryon de référence — et

nous affichons l'évolution du résidu moyen de la solution finale selon le stade de développement de l'embryon test.

La figure 5.16 présente des résultats d'appariements de cellules des images test (T1. 64), (T1. 112), (T2) et (T3) avec les images de références (Ref. 64) et (Ref. 112).

- Au stade 64 cellules, la segmentation de l'embryon de référence comporte 66 régions au lieu de 64 en raison de deux sur-segmentations. Dans la séquence temporelle (T1. 64) (respectivement (T2)), l'image que nous identifions au même stade de développement que (Ref. 64) possède 71 régions segmentées au lieu de 64, avec 7 cellules sur-segmentées (respectivement 69 régions segmentées avec 5 cellules sur-segmentées). Le recalage final entre (Ref. 64) et l'image identifiée de (T1. 64) (respectivement l'image identifiée de (T2)) permet de construire un total de 64 (respectivement 60) appariements de régions.
- Au stade 112 cellules, la segmentation de l'embryon de référence comporte 113 régions en raison de deux sur-segmentations et une sous-segmentation. Dans la séquence temporelle (T1. 112), l'image identifiée au même stade de développement que (Ref. 112) est segmentée en 133 régions au lieu de 112, avec 21 cellules sur-segmentées. L'image (T3) est quant à elle correctement segmentée en 114 régions. Le recalage final entre (Ref. 112) et l'image identifiée de (T1. 112) (respectivement (T3)) permet de construire un total de 106 (respectivement 103) appariements de régions.

Comme le montre la figure 5.16, où les images d'embryons test identifiées au même stade de développement que les images de référence sont indiquées par un nombre de cellules écrit en gras, l'inspection visuelle des appariements de régions obtenus par notre méthode de recalage permet d'établir que ceux-ci sont presque tous corrects. Cette figure permet aussi d'observer que lorsque le stade de développement de l'embryon test ne coïncide pas exactement avec celui de l'embryon de référence, les appariements de régions proposés par notre méthode restent globalement consistants — à l'exception de la première image de la séquence test (T1. 64), correspondant à un embryon au stade 43 cellules, et dont le recalage sur l'embryon (Ref. 64) n'est pas biologiquement consistant —, bien que l'hypothèse d'une bijection entre les correspondances de cellules des deux embryons ne soit plus valide dans cette condition.

La figure 5.17 illustre la fiabilité de la méthode d'extraction de la meilleure hypothèse d'appariements (solution globale) présentée en section 5.4.3. Elle présente, pour chacune des images test présentées en figure 5.16, la variation de la distance résiduelle entre les barycentres de régions appariées en fonction de l'initialisation de l'algorithme de recalage. L'extraction de la solution globale minimisant ce résidu moyen n'est fiable que si cette variation présente un minimum global manifeste. Dans le cas de l'image au stade 43 cellules de (T1. 64), mal recalée sur (Ref. 64), ainsi que de l'image de cette série au stade 54 cellules, on constate que la variation des résidus moyens selon l'initialisation du recalage ne fait pas apparaître de minimum franc. Il en va de même pour le recalage des images de (T1. 112) aux stades 76 et 99 cellules sur (Ref. 112). Pour les autres graphes présentés, reliés au recalage d'embryons à des stades de développement plus proches, on constate à l'inverse la présence d'un

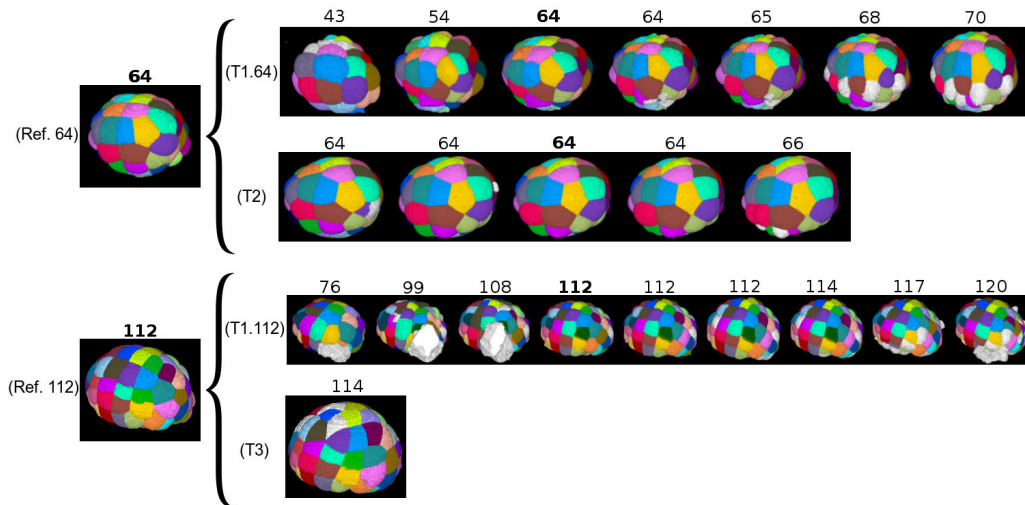


FIGURE 5.16 – Résultat du recalage des embryons test sur les images de référence. Les cellules appariées apparaissent de la même couleur. Les cellules test non appariées apparaissent en blanc. Les nombres situés au dessus des images d’embryons indiquent le nombre de cellules qui composent ces embryons. Les nombres écrits en gras indiquent les images des séries test que nous identifions comme étant au même stade de développement que les embryons de référence.

minimum bien marqué.

On observe ainsi que l’existence d’un minimum global franc est respectée lorsque le stade de développement des images test est le même que celui des images de référence, puis se détériore lorsque les stades de développement des embryons à recaler s’éloignent. La figure 5.18 illustre l’existence d’intervalles de « confiance », caractérisés par de petites distances résiduelles moyennes entre les centres de gravité des cellules appariées. Dans ces intervalles, le recalage proposé présente une fiabilité élevée. On observe que ces intervalles, loin d’être étroits, s’étalent sur plusieurs minutes avant et après les instants de développement de l’embryon test qui correspondent le mieux aux stades de développement de l’embryon de référence. Les intervalles de confiance estimés sur la figure 5.18 suggèrent qu’un embryon en développement peut être recalé sur un image d’embryon de référence avec un haut degré de fiabilité sur une fenêtre d’une douzaine de minutes. Nous montrons ainsi qu’avec la méthode de recalage d’embryons d’ascidies que nous présentons dans ce chapitre, il n’est pas indispensable que les stades de développement des embryons test et de référence que l’on souhaite recaler soient parfaitement alignés.

5.6 Conclusions et perspectives

Dans ce chapitre, nous avons présenté une méthode automatique de recalage spatial et d’identification des correspondances cellule à cellule entre des embryons d’ascidies à un même stade de développement. Cette méthode exploite le caractère

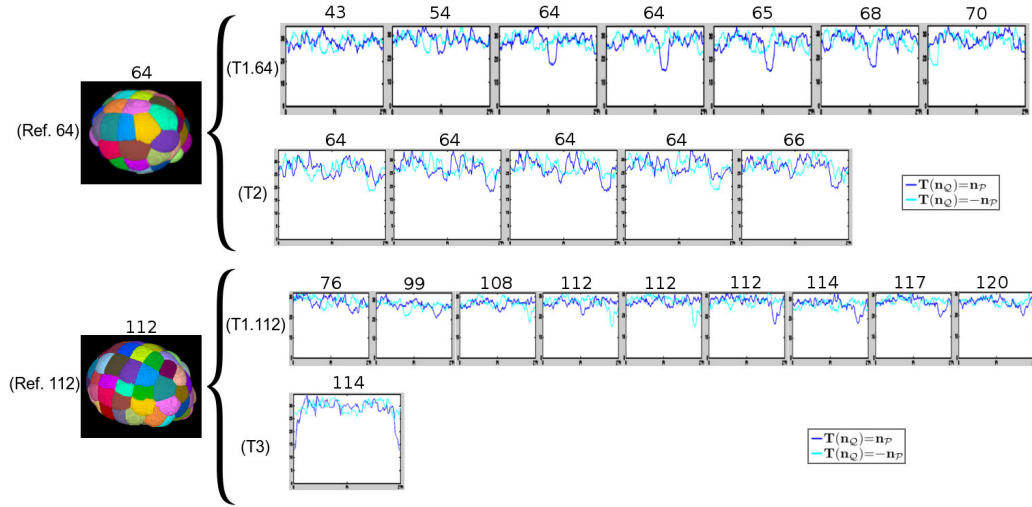


FIGURE 5.17 – Variation de la distance résiduelle moyenne tamisée entre les barycentres de cellules appariées en fonction de la transformation initiale $\mathbf{T} \in \mathcal{T}_0$ utilisée pour le recalage des images test de la figure 5.16 sur les images de référence. L'axe des abscisses représente l'angle de rotation appliqué à l'embryon test autour de son axe de symétrie lors de l'étape d'initialisation (section 5.4.1). L'axe des ordonnées représente la distance résiduelle moyenne tamisée exprimée en voxels. Les courbes bleues (respectivement cyan) expriment les variations de distance résiduelle moyenne selon $\mathbf{T} \in \mathcal{T}_0^+$ (respectivement $\mathbf{T} \in \mathcal{T}_0^-$). Pour chaque recalage, la transformation $\hat{\mathbf{T}}$ permettant d'obtenir les appariements illustrés en figure 5.16 est celle qui minimise le résidu moyen.

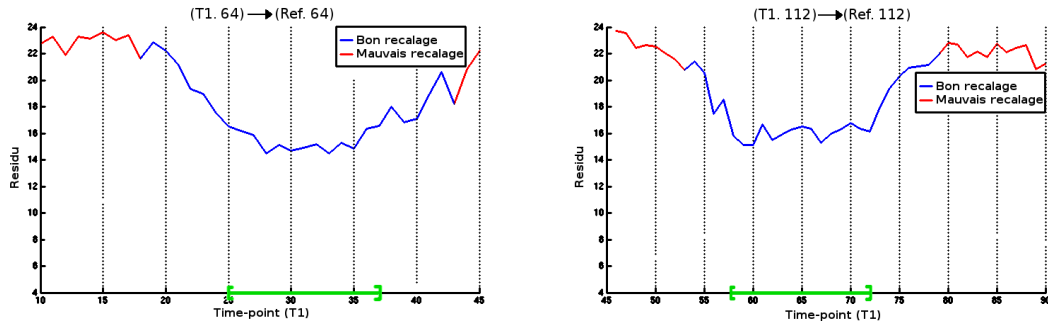


FIGURE 5.18 – Évolution du résidu moyen tamisé $\mathcal{C}_\alpha(\widehat{Map}, \widehat{\mathbf{T}})$ (équation 5.15) de l'embryon test (T1. 64) sur l'image de référence (Ref. 64) à gauche et de l'embryon test (T1. 112) sur l'image de référence (Ref. 112) à droite au cours du développement de l'embryon test. Une évaluation visuelle des résultats d'appariements des cellules nous permet de classer les résultats de recalage en deux ensembles selon leur pertinence biologique : les *bons* recalages (bleu) et les *mauvais* recalages (rouge). Les intervalles en vert indiquent des domaines de « confiance » à l'intérieur desquels les solutions globales (section 5.4.3) sont bien marquées.

fortement stéréotypé du développement des embryons d'ascidies et s'appuie sur l'extraction préalable de la symétrie bilatérale caractéristique de ces embryons jusqu'au stade neurula. Nous avons évalué la méthode proposée et démontré sa robustesse sur des données réelles provenant d'embryons de *Phallusia mammillata* marqués au niveau de leurs membranes cellulaires.

Des améliorations peuvent être apportées à ce travail pouvant être qualifié de préliminaire. Notre méthode de recalage n'établit en effet les hypothèses d'appariements de cellules que sur un critère de proximité des barycentres de cellules. Il est possible de considérer des critères supplémentaires afin de construire ces appariements, comme les adjacences de cellules ou l'exploitation du volume des cellules. Enfin, constituer un ensemble d'appariements de cellules peut être abordé sous l'angle de la théorie des graphes.

Il est possible d'envisager une adaptation de notre méthode de recalage d'embryons pour des problèmes de recalage d'images multimodales. L'annexe B présente une ébauche de travail réalisée dans ce sens. Dans cette annexe, nous investiguons l'exploitation de cette méthode pour recaler une image d'embryon marqué au niveau de ses noyaux de cellules sur une image d'embryon marqué au niveau de ses membranes.

Une autre perspective de ce travail est de l'étendre à un recalage temporel entre individus, permettant ainsi d'identifier des correspondances de stades de développement afin d'aligner des séquences temporelles d'images d'embryons.

La capacité d'identifier des correspondances de cellules entre différents embryons constitue une grande avancée pour l'embryologie puisque cela offre la possibilité de conduire des analyses statistiques portant sur des populations à l'échelle de la cellule. Un tel outil ouvre également la perspective de recaler un atlas (c'est-à-dire un embryon numérique modèle aux cellules parfaitement segmentées) sur n'importe quel embryon, ce qui rendrait par exemple possible la correction systématique d'inévitables erreurs de segmentation de cellules.

Recalage spatio-temporel de séquences 3D+ t de méristèmes floraux

Sommaire

6.1 Introduction	96
6.2 Augmentation de la résolution temporelle d'une séquence d'images 3D de méristème en développement	99
6.2.1 Compensation du déplacement rigide de la plante au fil des acquisitions	99
6.2.2 Recalage non linéaire entre paires d'images	100
6.2.3 Construction d'une séquence d'images interpolées	100
6.3 Recalage spatial de méristèmes floraux par une approche géométrique	103
6.3.1 Extraction de primitives géométriques	104
6.3.2 Recalage par la méthode du plus proche point itéré	104
6.3.3 Complexité algorithmique et choix d'implémentation	107
6.4 Recalage temporel d'une image sur une séquence	107
6.4.1 Extraction des primitives d'une image interpolée	108
6.4.2 Sélection de la meilleure correspondance temporelle	108
6.4.3 Critère de similarité	109
6.5 Stratégies de recalage de séquences de MF	113
6.5.1 Notations	113
6.5.2 Recalage image sur image	114
6.5.3 Recalage image sur séquence	115
6.5.4 Recalage séquence sur séquence	118
6.6 Expériences et résultats	119
6.6.1 Données	119
6.6.2 Interpolation d'images	121
6.6.3 Recalage spatial d'images de MF	123
6.6.4 Stratégies de recalage	130
6.7 Conclusions et perspectives	131

6.1 Introduction

En biologie développementale, l'analyse d'image constitue le principal moyen d'investigation pour quantifier à l'échelle de la cellule la croissance des tissus d'organismes en développement (animaux ou plantes) [Megason 2007]. Les techniques d'imagerie ont largement progressé ces dernières décennies, et il devient aujourd'hui possible d'étudier les mécanismes génétiques qui gouvernent le développement des embryons et des méristèmes (zones d'activité cellulaire chez les plantes) en imageant des séquences temporelles d'images 3D *in vivo* de ces organismes avec une excellente résolution spatiale et une résolution temporelle suffisamment fine pour suivre les déformations cellulaires au cours du temps [Keller 2013].

En plus du besoin d'outils d'analyse d'images et de séquences [Fernandez 2010a], l'étude de la morphogenèse fait également apparaître le besoin de comparer le développement d'individus, comme nous l'avons décrit dans la section 5.1. En ce sens, le développement de méthodes de recalage inter-individus constitue un enjeu de taille dans l'optique de réaliser des études quantitatives portant sur la variabilité de développement des organismes. Chez l'embryon, l'état de l'art propose plusieurs méthodes de recalage inter-individus exploitant un développement au caractère très stéréotypé de ce type d'organisme (voir la section 5.1). Le développement des organismes végétaux est cependant davantage complexe que celui des embryons d'ascidies, rendant plus difficile la mise au point d'outils de recalage inter-individus.

Le méristème apical caulinaire (MAC, *shoot apical meristem* en anglais) constitue la zone de croissance de la partie aérienne des plantes terrestres et est continuellement le siège de divisions cellulaires. Il est à l'origine de la formation de tous les organes situés au dessus du niveau du sol (fleurs, feuilles, tiges), ce qui en fait un objet d'intérêt majeur pour l'étude de l'organogenèse. Dans le cadre du projet de recherche Morphogenetics, nous nous intéressons au développement des organes floraux chez une plante modèle, *Arabidopsis thaliana*. Lorsque le processus de formation des fleurs débute, le méristème apical caulinaire évolue en méristème d'inflorescence duquel émergent un à un des méristèmes floraux selon une phyllotaxie spécifique régulée par les expressions géniques de la plante [Schultz 1991]. Dans ce chapitre, on considère le problème du recalage spatio-temporel entre une séquence de référence et une séquence test d'images 3D de méristèmes floraux (MF) d'*arabidopsis* sauvages en développement.

Le recalage spatial entre deux images de MF au même stade de développement consiste à déterminer une transformation géométrique optimale, au sens d'une fonction de coût, permettant de passer d'une image à l'autre dans un même référentiel. À l'inverse des embryons d'ascidies qui possèdent un développement très stéréotypé, avec l'existence de correspondances cellule à cellule entre les individus d'une même espèce, aucune correspondance équivalente entre cellules n'existe entre deux méristèmes floraux au même stade de développement. Un méristème se décompose en groupes de cellules, formant des couches (ou *layers*) distinctes (L1, L2, L3) où L1 et L2 constituent respectivement les couches épidermique et sous-épidermique tandis que L3 constitue le corpus du méristème (figure 6.2). Toutes

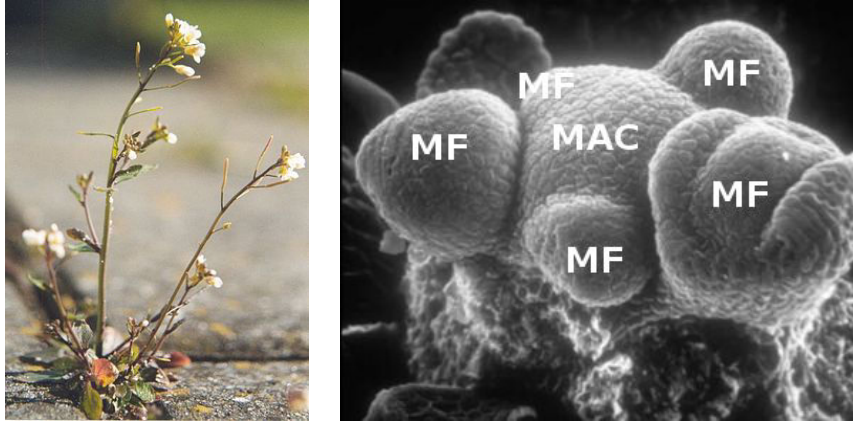


FIGURE 6.1 – Plant d'*Arabidopsis thaliana* et méristème d'inflorescence observé au microscope électronique à balayage. MAC : méristème apical caulinaire. MF : méristème floral.

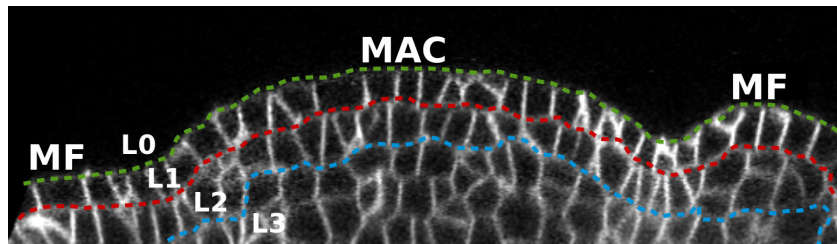


FIGURE 6.2 – Vue en coupe d'un méristème d'inflorescence et des interfaces séparant les couches de cellules L1, L2 et L3. Nous désignons par L0 l'extérieur du méristème.

les cellules de la couche L1 subissent des divisions anticlinales, c'est-à-dire perpendiculaires à la surface du méristème, de même que la majorité des cellules de la couche L2. Ainsi, les couches L1 et L2 restent distinctes durant toute la croissance du méristème [Barton 1993]. La forme globale des fleurs d'une même espèce est également très homogène [Legrand 2014], avec un stade de développement du MF identifiable à sa taille [Landrein 2014]. Considérant ces formes d'invariance entre méristèmes floraux d'*Arabidopsis*, un critère d'ajustement reposant sur les interfaces des couches cellulaires L1 et L2 peut être approprié pour construire une mesure de qualité de recalage entre individus à un même stade de développement. Les invariances sus-mentionnées laissent également supposer qu'il est possible de superposer deux images de méristèmes au même stade de développement à l'aide d'une transformation sans déformation (c'est-à-dire rigide). Nous ramenons ainsi le problème du recalage spatial d'images 3D de méristèmes à celui de recaler rigidement les interfaces entre couches de cellules.

Le problème du recalage temporel consiste à extraire pour une image 3D « flottante » de MF l'instant de développement auquel il s'identifie le mieux dans une séquence temporelle d'images de référence. Contrairement aux embryons d'asci-

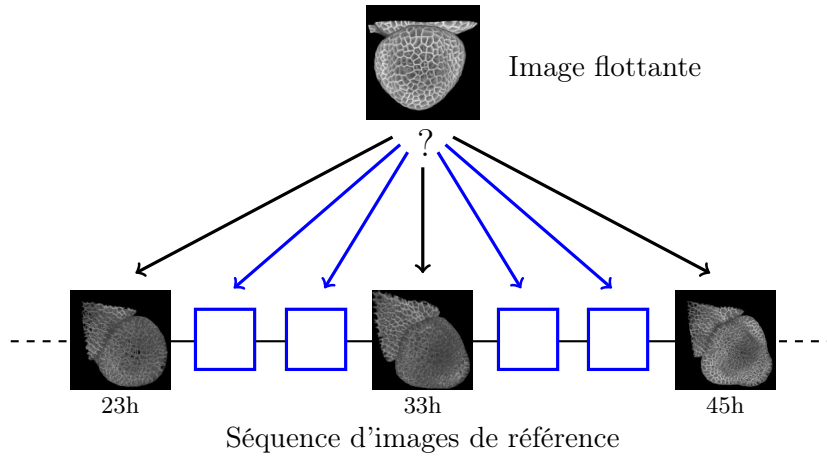


FIGURE 6.3 – Le problème de recalage spatio-temporel d’une image flottante de MF sur une séquence d’images de référence consiste à déterminer le recalage spatial de l’image flottante sur chacune des images de la série de référence, puis à déterminer l’image de référence sur laquelle l’image flottante se recalc de manière optimale. On observe ici une évolution importante de la morphologie du MF de référence entre deux images successives de la série. Afin d’assurer un recalage temporel précis de l’image flottante sur la séquence de référence, une étape préliminaire consiste donc à estimer des images intermédiaires (symbolisées en bleu sur la figure) entre les images successives de la séquence.

dies — imagés à l’aide d’un microscope à feuille de lumière assurant une excellente fréquence d’acquisition — les séquences temporelles d’images de MF sont obtenues par microscopie confocale avec un pas de temps allant de 5 à 13 heures entre deux acquisitions successives sur une durée de développement de la plante pouvant atteindre 5 à 6 jours. En raison du délai relativement important qui s’écoule entre deux acquisitions successives d’images de MF, la différence de stade de développement peut être conséquente entre une image de MF et celle qui lui est la plus proche dans une séquence. Cela rend l’identification de correspondances temporelles imprécise entre deux séquences temporelles d’images de MF. Une stratégie permettant une résolution plus fine de la correspondance temporelle entre une image et une séquence est d’augmenter artificiellement la résolution temporelle de la séquence en estimant des images intermédiaires (figure 6.3).

Ce chapitre se décompose en plusieurs parties. Dans un premier temps, nous proposons une méthodologie permettant de construire une séquence d’images à une résolution temporelle plus fine à partir d’une séquence originale par interpolation d’images (section 6.2). Puis nous présentons un outil de recalage spatial de MF (section 6.3). Nous proposons ensuite une stratégie de recalage temporel entre une image test et une séquence — à la résolution temporelle préalablement augmentée — de référence (section 6.4). Enfin, nous présentons une stratégie de recalage entre

deux séquences temporelles d’images de MF (section 6.5). La section 6.6 présente les expériences réalisées sur les méthodes que nous proposons. Finalement, nous concluons sur notre travail en section 6.7.

6.2 Augmentation de la résolution temporelle d’une séquence d’images 3D de méristème en développement

L’augmentation de la résolution temporelle d’une série d’images peut être réalisée en estimant des images intermédiaires entre deux acquisitions successives par interpolation. Dans cette section, nous présentons une méthode d’interpolation d’images permettant de construire une image $I(t)$ à partir d’une séquence temporelle d’images $\{I_{t_i}\}_{i \in \{0:n\}}$ pour toute valeur de $t \in [t_0, t_n]$.

6.2.1 Compensation du déplacement rigide de la plante au fil des acquisitions

Entre deux acquisitions successives d’images I_{t_i} et $I_{t_{i+1}}$, un réajustement du champ de vue du microscope est nécessaire pour suivre les déplacements rigides du MF en développement en raison de la croissance de la plante observée. De façon à compenser les déplacements du MF dans le champ de vue entre deux acquisitions, nous procédons au calcul de la transformation $\mathbf{R}_{t_{i+1} \leftarrow t_i}$ recalant rigidement I_{t_i} sur $I_{t_{i+1}}$.

Le recalage rigide entre les images successives de la série est réalisé à l’aide de la méthode hybride de recalage par blocs (*block matching*) [Ourselin 2000]. En reprenant les notions de recalage introduites dans la section 5.2, le principe de base de cette méthode est de découper les images à recaler en petits blocs, de considérer chaque bloc comme une primitive iconique et d’appliquer sur ces primitives un schéma itératif d’optimisation analogue à la méthode des plus proches points itérés [Besl 1992]. Les appariements de primitives sont construits en maximisant une métrique de similarité entre les blocs d’images.

Le choix d’un champ de vue commun à toutes les images de la série se fait à partir de l’ensemble des transformations rigides $\{\mathbf{R}_{t_{i+1} \leftarrow t_i}\}_{i \in \{0:n-1\}}$. On ré-échantillonne finalement l’ensemble des images de la série dans ce champ de vue en leur appliquant une combinaison des transformations calculées. Ainsi, nous avons construit une nouvelle série d’images dans laquelle nous avons compensé les déplacements rigides inhérents au protocole d’acquisition des données.

Dans la suite de la section, afin d’éviter l’alourdissement inutile des notations, nous désignons par $\{I_{t_i}\}_{i \in \{0:n\}}$ la séquence d’images de MF dont les déplacements rigides sont déjà compensés.

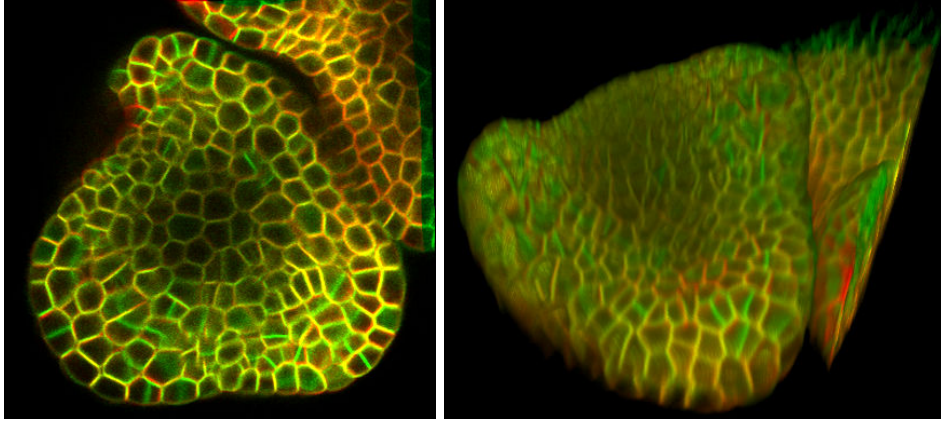


FIGURE 6.4 – Vue en coupe (gauche) et vue 3D (droite) du résultat du recalage non linéaire entre une deux images consécutives d’un MF. En rouge : $I_{t_i} \circ \mathbf{T}_{t_i \leftarrow t_{i+1}}$. En vert : $I_{t_{i+1}}$.

6.2.2 Recalage non linéaire entre paires d’images

Nous considérons le problème du recalage entre deux images I_{t_i} et $I_{t_{i+1}}$ d’un MF. Les déformations subies par le MF au cours de son développement étant non linéaires, nous cherchons une transformation sous forme de champ de déformations $\mathbf{T}_{t_i \leftarrow t_{i+1}} = \text{Id} + \mathbf{u}_{t_i \leftarrow t_{i+1}}$:

$$\forall M \in \mathcal{D}(I_{t_{i+1}}), \quad \mathbf{T}_{t_i \leftarrow t_{i+1}}(M) = M + \mathbf{u}_{t_i \leftarrow t_{i+1}}(M), \quad (6.1)$$

où $\mathcal{D}(I)$ est le domaine de définition d’une image I .

Cette transformation est calculée en deux temps. Nous procédons d’abord à un recalage linéaire (affine) entre les deux images — que l’on suppose initialement rigidement recalées d’après la section 6.2.1. La transformation affine trouvée est ensuite utilisée comme une initialisation pour le recalage non linéaire (figure 6.4).

Les recalages affine et non linéaire sont réalisés à l’aide de la méthode de recalage par blocs [Ourselin 2000].

Note. Entre deux acquisitions d’images, il est fréquent que des variations de contraste surviennent (figures 6.5a et 6.5b). Ces variations de contraste peuvent être estimées lorsque les images I_{t_i} et $I_{t_{i+1}}$ sont recalées non linéairement en réalisant une ajustement de contraste par égalisation d’histogramme (figures 6.5c et 6.5d). Nous considérons dans la suite que les contrastes des images de la séquence considérée sont ajustés deux à deux.

6.2.3 Construction d’une séquence d’images interpolées

Connaissant la transformation non linéaire $\mathbf{T}_{t_i \leftarrow t_{i+1}}$ permettant de projeter I_{t_i} sur $I_{t_{i+1}}$, nous cherchons à estimer les transformations intermédiaires $\mathbf{T}_{t_i \leftarrow t}$ et $\mathbf{T}_{t_{i+1} \leftarrow t}$ projetant respectivement I_{t_i} et $I_{t_{i+1}}$ à l’instant de développement $t \in [t_i, t_{i+1}]$ afin de construire l’image interpolée $I(t)$.

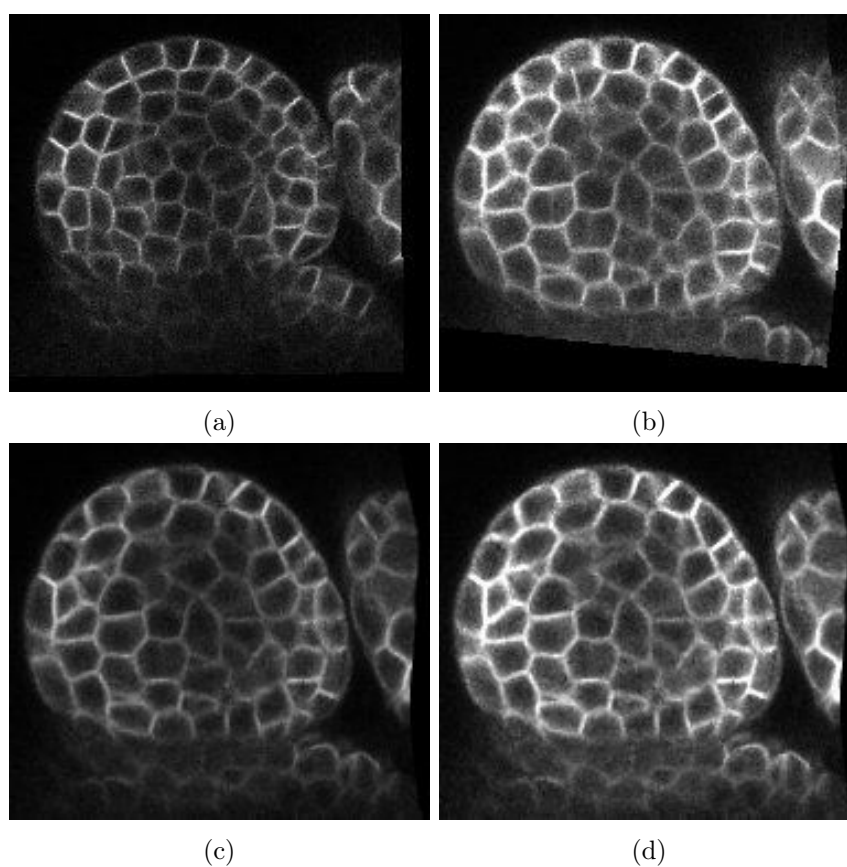


FIGURE 6.5 – Égalisation de contraste. Vues en coupe de : (a) l'image I_{t_i} ; (b) l'image $I_{t_{i+1}}$; (c) l'image projetée $I_{t_i} \circ \mathbf{T}_{t_i \leftarrow t_{i+1}}$; (d) le résultat de l'ajustement du contraste de l'image projetée $I_{t_i} \circ \mathbf{T}_{t_i \leftarrow t_{i+1}}$ par rapport à l'image $I_{t_{i+1}}$.

6.2.3.1 Transformations intermédiaires

Pour estimer les transformations intermédiaires, nous supposons que chaque point composant le MF observé se déplace linéairement et à vitesse constante de sa position au temps t_i à sa position au temps t_{i+1} (méthode d'interpolation *Intim*¹ [Rey 2002, annexe E]). Ainsi, en posant le coefficient d'interpolation $\gamma = \frac{t-t_i}{t_{i+1}-t_i}$, on a :

$$\forall M \in \mathcal{D}(I_{t_{i+1}}), \quad \mathbf{T}_{t \leftarrow t_{i+1}}(M) = M + (1 - \gamma)\mathbf{u}_{t_i \leftarrow t_{i+1}}. \quad (6.2)$$

On en déduit donc :

$$\forall t \in [t_i, t_{i+1}], \quad \mathbf{u}_{t \leftarrow t_{i+1}} = (1 - \gamma)\mathbf{u}_{t_i \leftarrow t_{i+1}}. \quad (6.3)$$

Finalement, la transformation $\mathbf{T}_{t_{i+1} \leftarrow t}$ est calculée comme étant l'inverse de la transformation $\mathbf{T}_{t \leftarrow t_{i+1}} = \text{Id} + \mathbf{u}_{t \leftarrow t_{i+1}}$. La supposition de la linéarité du déplacement des points entre t_i et t_{i+1} permet de déduire directement que les champs de vecteurs $\mathbf{u}_{t_i \leftarrow t}$ et $\mathbf{u}_{t_{i+1} \leftarrow t}$ sont proportionnels et vérifient la relation :

$$\forall t \in]t_i, t_{i+1}[, \quad \mathbf{u}_{t_i \leftarrow t} = -\frac{\gamma}{1 - \gamma}\mathbf{u}_{t_{i+1} \leftarrow t}, \quad (6.4)$$

menant à $\mathbf{T}_{t_i \leftarrow t} = \text{Id} + \mathbf{u}_{t_i \leftarrow t}$.

6.2.3.2 Interpolation d'images en niveaux de gris

L'image interpolée $I(t)$ s'obtient en combinant les images d'intensité ré-échantillonnées $I_{t_i} \circ \mathbf{T}_{t_i \leftarrow t}$ et $I_{t_{i+1}} \circ \mathbf{T}_{t_{i+1} \leftarrow t}$:

$$I(t) = (1 - \gamma)I_{t_i} \circ \mathbf{T}_{t_i \leftarrow t} + \gamma I_{t_{i+1}} \circ \mathbf{T}_{t_{i+1} \leftarrow t}, \quad (6.5)$$

où γ est le coefficient d'interpolation introduit en début de section. La figure 6.6 présente un exemple d'interpolation d'image.

Note. Le coefficient d'interpolation γ est le même sur l'ensemble du domaine de définition de $I(t)$. Cela signifie en particulier que les divisions cellulaires du MF entre deux instants t_i et t_{i+1} apparaissent de manière synchronisée partout dans l'image, ce qui est probablement irréaliste. Des études mettent en évidence le caractère stéréotypé des divisions cellulaires chez le méristème apical caulinaire de l'arabidopsis [Shapiro 2015], laissant penser qu'il est possible de prédire l'instant auquel survient une division cellulaire. La prise en compte d'un éventuel modèle prédictif pour l'occurrence des divisions cellulaires peut être intégrée dans notre méthode d'interpolation d'images en considérant un coefficient d'interpolation calculé localement au regard du modèle.

1. *Intim* : INTerpolation of IMages.

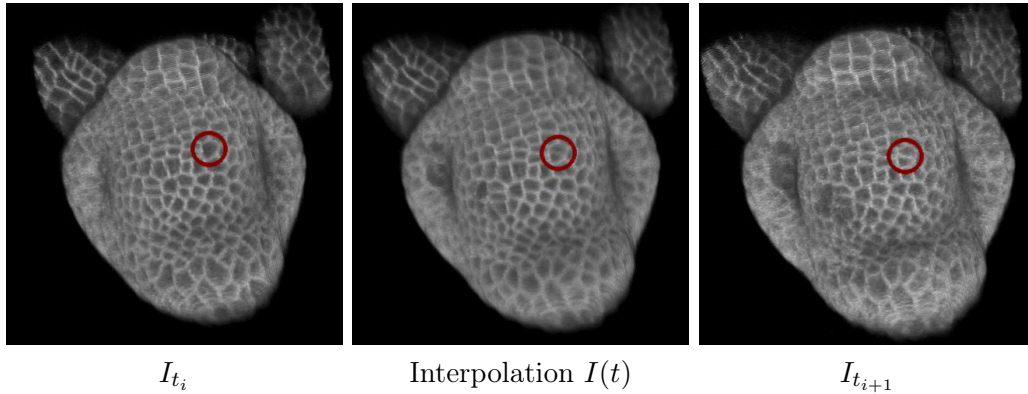


FIGURE 6.6 – Résultat de l’interpolation de deux images de MF pour $t = \frac{t_i + t_{i+1}}{2}$ ($\gamma = 0,5$). Le cercle rouge met en évidence une division cellulaire.

6.2.3.3 Interpolation d’images segmentées

Le calcul des transformations intermédiaires $\mathbf{T}_{t_i \leftarrow t}$ et $\mathbf{T}_{t_{i+1} \leftarrow t}$ permet de ré-échantillonner les images segmentées provenant des instants t_i et t_{i+1} à tout instant intermédiaire $t \in]t_i, t_{i+1}[$. Cependant, il n’existe aucune garantie pour que les segmentations ainsi transformées soient compatibles entre elles, au sens où les frontières formées par les groupes de labels (cellules) qui se correspondent se superposent parfaitement.

La détermination de correspondances de labels entre les images segmentées aux instants t_i et t_{i+1} à l’aide de l’outil de comparaison de segmentation présenté au chapitre 3, conjointement à la connaissance des transformations intermédiaires $\mathbf{T}_{t_i \leftarrow t}$ et $\mathbf{T}_{t_{i+1} \leftarrow t}$, permet d’estimer pour tout instant $t \in]t_i, t_{i+1}[$ des segmentations intermédiaires provenant des instants t_i et t_{i+1} compatibles entre elles, c’est-à-dire parfaitement superposables. L’outil d’interpolation de segmentations multi-labels que nous avons construit est une extension des méthodes d’interpolation basées sur la forme (*Shape-based interpolation*) [Herman 1992]. Nous décrivons cet outil, permettant de rendre « compatibles » des images segmentées intermédiaires, dans l’annexe C.

6.3 Recalage spatial de méristèmes floraux par une approche géométrique

Dans cette section, nous considérons le problème du recalage spatial entre deux méristèmes floraux I (référence) et J (test) à un même stade de développement. Les invariances de taille [Landrein 2014] et de forme [Legrand 2014] entre individus à un même stade de développement nous amènent à considérer un recalage rigide entre I et J (c’est-à-dire sans déformation). De plus, le recalage de I et J peut être ramené à un problème de recalage des couches L1 et L2 en raison de la conservation des couches de cellules L1 et L2 au cours du développement d’un MF [Barton 1993]. Dans la suite

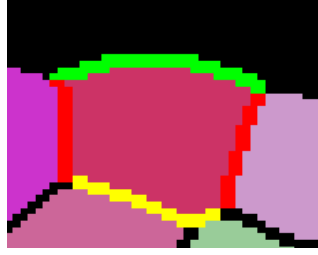


FIGURE 6.7 – L’interface d’une cellule appartenant à la couche L1 se subdivise en l’interface fond/cellule (en vert), l’interface cellule/L1 (en rouge) et l’interface cellule/L2 (en jaune). Les cellules de L1 sont caractérisées par leur contact avec le fond (L0). Les cellules de L2 sont caractérisées par leur contact avec la couche L1.

de cette section, nous proposons d’extraire des primitives géométriques à partir des images à recaler, puis de les mettre en correspondance à l’aide de l’algorithme du plus proche point itéré.

6.3.1 Extraction de primitives géométriques

Nous considérons dans cette partie que les segmentations de cellules des images I et J sont connues — se référer à [Fernandez 2010a] ou au chapitre 4 pour des méthodes de segmentation de cellules. Les cellules qui composent la couche L1 d’un MF sont aisément identifiables comme étant en contact direct avec l’extérieur du méristème (que nous notons L0), c’est-à-dire le label de fond de l’image segmentée (figure 6.7). De la même façon, on identifie les cellules de la couche L2 comme étant celles en contact avec la couche L1. Nous définissons les notations suivantes :

- $k = 1$ (respectivement 2, 3) désigne l’indice correspondant à l’interface L0-L1 (respectivement l’interface L1-L2, L2-L3) ;
- K désigne un ensemble d’indices d’interfaces (par exemple $K = \{1, 2\}$) ;
- S_k^I (respectivement S_k^J) est l’interface correspondant à l’indice k de l’image I (respectivement J) ;
- S_K^I (respectivement S_K^J) est l’ensemble d’interfaces $\{S_k^I\}_{k \in K}$ (respectivement $\{S_k^J\}_{k \in K}$) ;
- P_k^I (respectivement P_k^J) est un ensemble de points constituant un échantillonnage de S_k^I (respectivement S_k^J) ;
- P_K^I (respectivement P_K^J) est l’ensemble d’ensembles de points $\{P_k^I\}_{k \in K}$ (respectivement $\{P_k^J\}_{k \in K}$).

La figure 6.8 présente le résultat de l’extraction des ensembles de points P_K^I situés aux interfaces S_K^I d’une image I de MF pour $K = \{1, 2\}$.

6.3.2 Recalage par la méthode du plus proche point itéré

Nous considérons ici un ensemble de points test P_K^J à recaler sur un ensemble de points de référence P_K^I . Le recalage spatial entre ces deux ensembles consiste à déterminer la transformation optimale $\hat{\mathbf{T}}_{I \leftarrow J}$ parmi l’ensemble des transformations

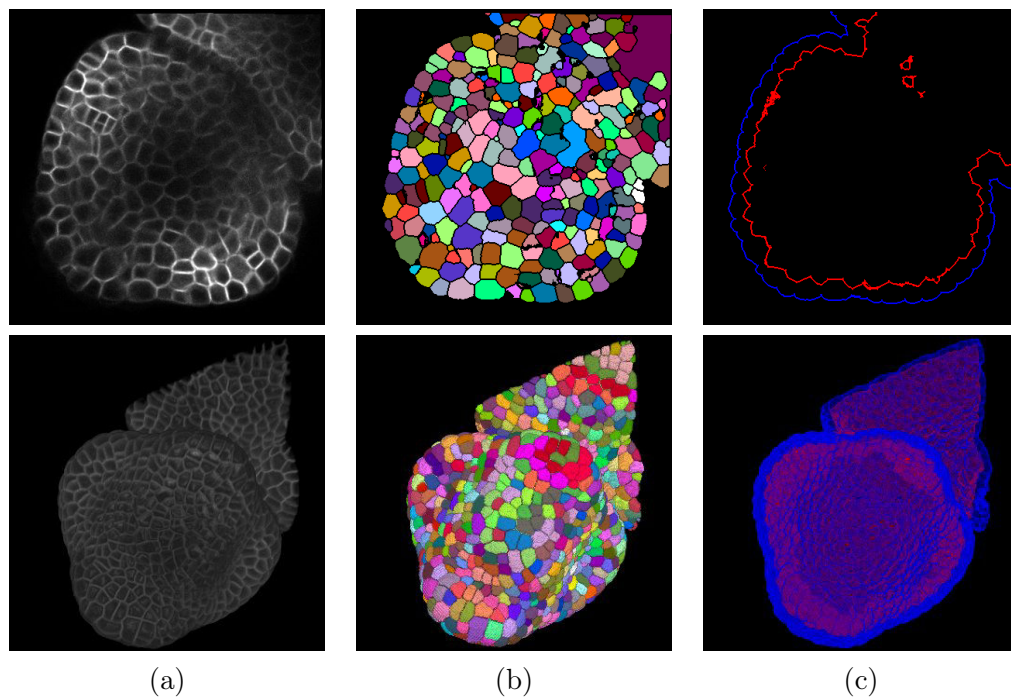


FIGURE 6.8 – Extraction des ensembles de points P_K^I à l'interface entre les couches de cellules L0-L1 et L1-L2 d'une image I . Première ligne : vues en coupe. Deuxième ligne : vues 3D. (a) Image d'intensité I . (b) Segmentation des cellules de l'image. (c) Points extraits à l'interface $S_{k=1}^I$ entre L0 et L1 (bleu) et $S_{k=2}^I$ (rouge) entre L1 et L2.

rigides $\text{Rig}(\mathbb{R}^3)$ recalant P_K^J sur P_K^I au sens d'une fonction de coût \mathcal{C} :

$$\widehat{\mathbf{T}}_{I \leftarrow J} = \arg \min_{\mathbf{T} \in \text{Rig}(\mathbb{R}^3)} \mathcal{C}(P_K^I, \mathbf{T}(P_K^J)). \quad (6.6)$$

Pour réaliser cette tâche, nous employons la méthode du plus proche point itéré [Besl 1992] — présenté en section 5.2. Ce procédé permet de minimiser itérativement la distance entre deux nuages de points. À chaque itération i , on calcule la meilleure transformation rigide $\mathbf{T}^{(i)}$ superposant un ensemble transitoire de points appariés $Map^{(i-1)}$ au sens des moindres carrés tamisés [Rousseeuw 1987] (équation 6.7).

$$\mathbf{T}^{(i)} = \arg \min_{\mathbf{T} \in \text{Rig}(\mathbb{R}^3)} \mathcal{C}_\alpha \left(Map^{(i-1)}, \mathbf{T} \right), \quad (6.7)$$

où la fonction de coût \mathcal{C}_α est définie par l'équation 5.15 (section 5.4.2). Le choix d'un estimateur aux moindres carrés tamisés permet d'exclure du calcul de chaque transformation $\mathbf{T}^{(i)}$ les appariements les plus aberrants en choisissant la proportion d'appariements conservés $p_\alpha \in]0, 1]$.

Un nouvel ensemble d'appariements de points $Map^{(i)} = Map(\mathbf{T}^{(i)})$ entre les ensembles P_K^I et P_K^J est calculé à l'itération i en recherchant pour chaque point $q \in P_K^J$ son plus proche voisin dans P_K^I appartenant à la même interface d'indice $k \in K$ (équation 6.8) :

$$\forall (p, q) \in P_K^I \times P_K^J, \quad (p, q) \in Map(\mathbf{T}) \Leftrightarrow \exists k \in K, \quad \begin{cases} q \in P_k^J \\ p = \arg \min_{\tilde{p} \in P_k^I} \text{dist}(\tilde{p}, \mathbf{T}(q)) \end{cases} \cdot \quad (6.8)$$

Nous résumons ici les étapes de l'algorithme du plus proche point itéré :

1. Initialisation de la transformation $\mathbf{T}^{(0)} := \mathbf{T}_{I \leftarrow J}^{\text{init}}$.
2. Calcul des appariements de points $Map^{(0)} = Map(\mathbf{T}^{(0)})$.
3. $i \leftarrow 0$.
4. Procéder jusqu'à convergence de $\{\mathbf{T}^{(i)}\}_{i \geq 0}$ à :
 - (a) $i \leftarrow i + 1$.
 - (b) Calcul de la transformation rigide $\mathbf{T}^{(i)}$ optimisant le recalage de P_K^J sur P_K^I (équation 6.7).
 - (c) Calcul des appariements de points $Map^{(i)} = Map(\mathbf{T}^{(i)})$ (équation 6.8).
5. On note $\widehat{\mathbf{T}}_{I \leftarrow J} := \mathbf{T}^{(i)}$ et $\widehat{Map}_{I \leftarrow J} := Map(\widehat{\mathbf{T}}_{I \leftarrow J})$.

Nous considérons que la série de transformations $\{\mathbf{T}^{(i)}\}_{i \geq 0}$ a convergé lorsqu'aucun point appartenant au domaine de définition $\mathcal{D}(J)$ de l'image test J ne se déplace d'une distance supérieure à la résolution spatiale d'un voxel entre deux transformations successives.

Le procédé débute par une étape d'initialisation où l'on détermine manuellement une transformation $\mathbf{T}_{I \leftarrow J}^{\text{init}}$ recalant approximativement P_K^J sur P_K^I . Cette étape peut être réalisée à l'aide d'un logiciel permettant la visualisation synchronisée d'objets 3D (par exemple Fiji [Schindelin 2012]).

6.3.3 Complexité algorithmique et choix d'implémentation

L'étape la plus chronophage de l'algorithme de recalage spatial que nous proposons est l'étape de calcul des appariements de points $Map^{(i)}$ répétée à chaque itération. La complexité de cette étape est proportionnelle au nombre de points à appairer, c'est-à-dire au cardinal de l'ensemble P_K^J (noté $\#P_K^J$), et proportionnelle au coût du processus de recherche du plus proche voisin d'un point p de P_K^J dans P_K^I (pour p appartenant à l'interface d'indice $k \in K$, ce coût dépend de $\#P_k^I$).

Le facteur de coût $\#P_K^J$ peut être contrôlé en réalisant un sous-échantillonnage de l'ensemble des interfaces test S_K^J à l'aide d'une méthode de parcellisation de surfaces [Flandin 2002]. Ensuite, nous avons privilégié une représentation des ensembles de points P_K^I sous forme d'arbres binaires de recherche k -d [Bentley 1975]. Cette structure de données permet de stocker des points et le coût de la recherche du plus proche voisin d'un point dans un arbre k -d est logarithmique (contre un coût linéaire pour une représentation des points sous forme de liste). Ainsi, l'étape d'appariements de points possède une complexité en $O(\#P_K^J \ln(\#P_k^I))$.

Typiquement, si un ensemble d'interfaces test S_K^J est parcellisé en 10^5 points et si chaque nuage de points P_k^I est représenté sous forme d'arbre k -d comptant $2,5 \cdot 10^5$ points, chaque itération de l'algorithme du plus proche point itéré s'exécute en moins d'un demi-dixième de seconde. Le nombre d'itérations nécessaires à la convergence de l'algorithme dépend de l'écart entre l'initialisation du processus et la solution optimale. L'expérience montre que la convergence est atteinte généralement en moins de 200 itérations, soit en moins de 10 secondes.

6.4 Recalage temporel d'une image sur une séquence

Nous considérons dans cette section le problème du recalage spatio-temporel d'une image test J de MF sur une séquence d'images de référence $\{I_{t_i}\}_{i \in \{0:n\}}$ provenant d'un MF de référence dont l'intervalle d'observation englobe le stade de développement correspondant à l'image J . Le but de cette section est de déterminer l'instant de développement $\hat{t}(J) \in [t_0, t_n]$ du MF de référence se rapprochant le plus du stade de développement de J .

Le procédé d'interpolation d'images proposé en section 6.2.3 permet de construire à partir de la séquence d'origine une séquence d'images interpolées $\{I(t)\}_{t \in T}$ à partir d'un échantillonnage $T = \{t_{i,\delta}\}_{i \in \{0:n\}, \delta}$ de l'intervalle de temps $[t_0, t_n]$. La séquence est supposée interpolée à intervalles de temps réguliers, c'est-à-dire $t_{i,\delta} = t_i + \delta \Delta t$, où le pas de temps Δt choisi permet d'assurer une faible variation morphologique du MF de référence entre deux instants consécutifs (typiquement, $\Delta t = 1\text{h}$).

Pour déterminer l'instant de développement $\hat{t}(J)$ du MF de référence s'identifiant au mieux au stade de développement de l'image J , nous proposons de recalculer spatialement J sur chacune des images de $\{I(t)\}_{t \in T}$ et d'extraire l'instant de T offrant le « meilleur » recalage (figure 6.9). Il s'agit pour cela de définir un critère de similarité qui sera maximisé pour $t = \hat{t}(J)$.

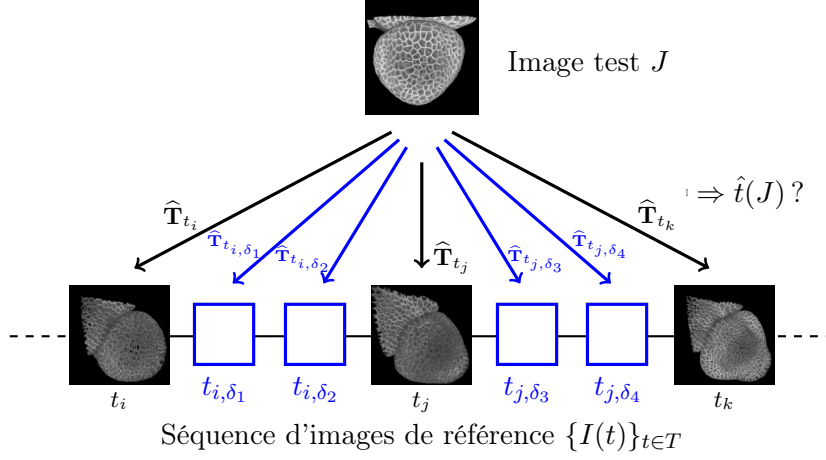


FIGURE 6.9 – La méthode de recalage temporel d’une image test J sur une séquence interpolée de référence $\{I(t)\}_{t \in T}$ consiste à calculer la transformation $\widehat{\mathbf{T}}_{I(t) \leftarrow J}$ (notée ici $\widehat{\mathbf{T}}_t$ pour la lisibilité du schéma) réalisant le recalage spatial de J sur chacune des images de la séquence de référence, puis à extraire l’instant de développement $\hat{t}(J) \in T$ offrant le meilleur recalage spatial de J sur la séquence.

6.4.1 Extraction des primitives d’une image interpolée

La méthode d’interpolation d’images segmentées présentée en section 6.2.3.3 et détaillée en annexe C permet de construire pour tout instant $t \in [t_0, t_n]$ une estimée de l’image des cellules segmentées du MF considéré.

Nous appliquons donc la méthode d’extraction des interfaces de couches de cellules de MF présentée en section 6.3.1 à la séquence interpolée des images de cellules segmentées aux instants $t \in T$. Nous construisons ainsi la séquence de surfaces $\{S_K^I(t)\}_{t \in T}$ (vérifiant pour tout $i \in \{0 : n\}$, $S_K^I(t_i) = S_K^{I t_i}$).

Nous construisons ainsi la séquence $\{P_K^I(t)\}_{t \in T}$ des ensembles de primitives que l’on obtient à partir de l’échantillonnage de la séquence des ensembles de surfaces $\{S_K^I(t)\}_{t \in T}$.

6.4.2 Sélection de la meilleure correspondance temporelle

Le problème du recalage temporel d’une image J sur une séquence interpolée $\{I(t)\}_{t \in T}$ revient à déterminer l’instant de développement $\hat{t}(J)$ qui optimise un critère de similarité s entre les ensembles de primitives P_K^J et $\{P_K^I(t)\}_{t \in T}$, auparavant spatialement recalés (i.e. $\widehat{\mathbf{T}}_{I(t) \leftarrow J}$ est connu pour tout $t \in T$, voir figure 6.10) :

$$\hat{t}(J) = \arg \max_{t \in T} s(t). \quad (6.9)$$

Nous souhaitons attirer l’attention du lecteur sur le fait que le calcul des transformations optimales $\widehat{\mathbf{T}}_{I(t) \leftarrow J}$ par la méthode proposée en section 6.3 nécessite la détermination manuelle d’une transformée initiale $\mathbf{T}_{I(t) \leftarrow J}^{\text{init}}$ pour chaque $t \in T$. Cepen-

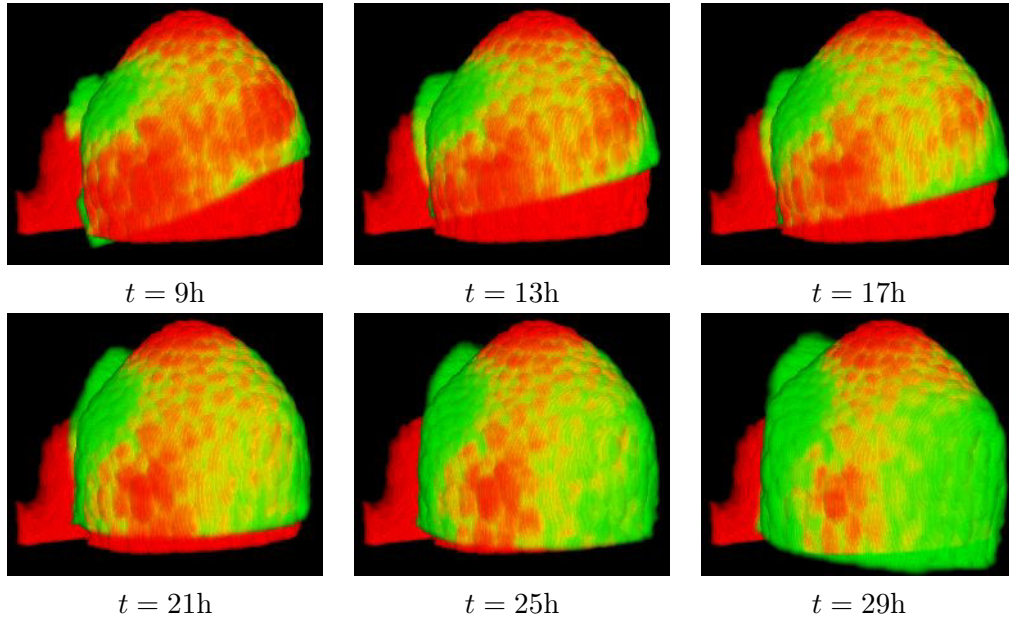


FIGURE 6.10 – Visualisation du recalage spatial d'une image test J (rouge) sur différentes images de la séquence de référence $\{I(t)\}_{t \in T}$ (vert).

dant, considérant que les mouvements rigides du MF observé sont compensés dans la séquence interpolée $\{I(t)\}_{t \in T}$ (se référer à la section 6.2.1), des stratégies de « propagation » de l'initialisation du recalage spatial d'une image sur une séquence peuvent être envisagées. Ainsi, une détermination manuelle d'une transformation $\mathbf{T}_{I(t) \leftarrow J}^{\text{init}}$ ne serait nécessaire que pour un seul $t \in T$, les transformations initiales pour les autres valeurs de t pouvant être calculées par propagation du recalage spatial. Dans la section 6.5, nous présentons différentes stratégies de propagation de l'initialisation du recalage d'une image sur une séquence.

6.4.3 Critère de similarité

Le recalage temporel de J sur la séquence $\{I(t)\}_{t \in T}$ doit permettre d'extraire l'instant $\hat{t}(J) \in T$ pour lequel la transformation de l'ensemble de points P_K^J s'aligne au mieux à l'ensemble de points cible $P_K^I(\hat{t}(J))$. Meilleur le recalage spatial, plus les distances résiduelles entre points appariés doivent être petites.

Nous proposons d'utiliser un critère de similarité basé sur l'histogramme cumulé des distances résiduelles entre points appariés après convergence du recalage. Soit la fonction $h_{I(t) \leftarrow J}$ d'histogramme cumulé des distances résiduelles entre points appariés après recalage :

$$h_{I(t) \leftarrow J} : d \in \mathbb{R}_+ \mapsto \# \left\{ (p, q) \in \widehat{Map}_{I(t) \leftarrow J} : \text{dist} \left(p, \widehat{\mathbf{T}}_{I(t) \leftarrow J}(q) \right) \leq d \right\}. \quad (6.10)$$

Ainsi, pour une distance d fixée, on peut définir un instant optimal $\hat{t}_d(J)$ maximisant l'histogramme cumulé des distances résiduelles au cours du temps (fig-

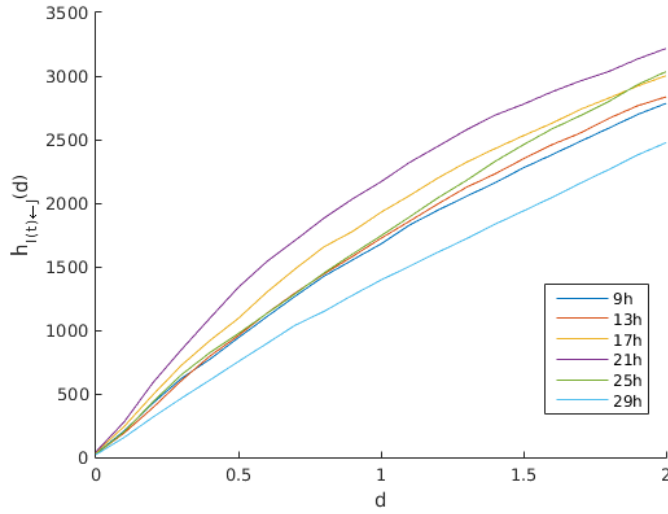


FIGURE 6.11 – Histogramme cumulé des distances résiduelles $h_{I(t)←J}(d)$ après recalage de l'image test J sur différentes images de la séquence $\{I(t)\}_{t∈T}$ (voir la figure 6.10). Les valeurs prises par l'histogramme cumulé sont les plus grandes pour $t = 21h$, traduisant un alignement temporel probable de l'image test sur cet instant de développement de la séquence de référence.

ure 6.11) :

$$\hat{t}_d(J) = \arg \max_{t∈T} h_{I(t)←J}(d). \quad (6.11)$$

Plutôt que de devoir fixer de façon arbitraire le paramètre de distance d à partir duquel on extrait l'instant $\hat{t}_d(J)$ maximisant $t \mapsto h_{I(t)←J}(d)$, nous souhaitons construire un critère de similarité exploitant plusieurs valeurs de d prises dans un intervalle discrétisé $D = \{d_0, \dots, d_m\}$ (l'instant à valeur maximale $\hat{t}_d(J)$ pouvant être variable selon $d \in D$). Intuitivement, on cherche à identifier un instant $\hat{t}(J)$ qui maximise de manière « consensuelle » pour les différentes valeurs de $d \in D$ l'histogramme cumulé des distances résiduelles $t \mapsto h_{I(t)←J}(d)$. Afin de comparer les histogrammes cumulés calculés pour différents d , nous proposons de normaliser pour chaque valeur de $d \in D$ la fonction $t \mapsto h_{I(t)←J}(d)$ par son maximum atteint en $\hat{t}_d(J)$. Nous définissons ainsi la mesure d'histogramme cumulé normalisé des distances résiduelles $n_{I(t)←J}$ illustrée par les figures 6.12 et 6.13 :

$$n_{I(t)←J} : d \in \mathbb{R}_+ \mapsto \frac{h_{I(t)←J}(d)}{\max_{\hat{t} \in T} h_{I(\hat{t})←J}(d)}. \quad (6.12)$$

Cette mesure vérifie à d fixé que l'instant $\hat{t}_d(J)$ défini par l'équation 6.11 maximise la fonction $t \mapsto n_{I(t)←J}(d)$, avec $n_{I(\hat{t}_d(J))←J}(d) = 1$.

Pour extraire l'instant $\hat{t}(J)$ maximisant de manière « consensuelle » la mesure normalisée $n_{I(t)←J}$, nous définissons finalement la fonction de similarité s de l'équation 6.9, maximisée par définition en $t = \hat{t}(J)$, par l'intégration de $d \mapsto n_{I(t)←J}(d)$

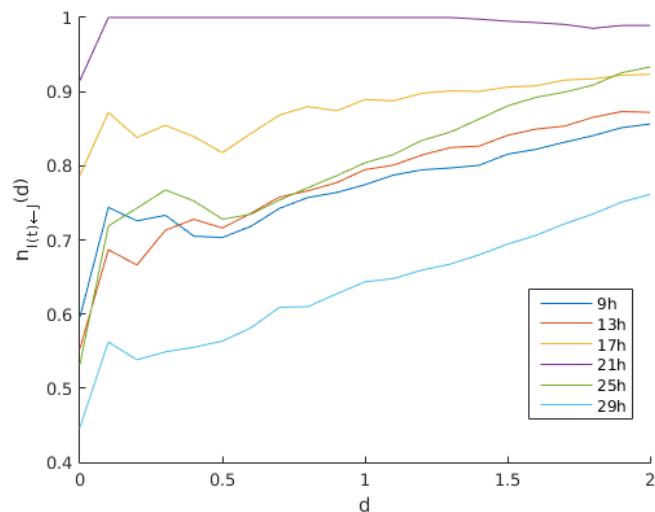


FIGURE 6.12 – Normalisation $n_{I(t) \leftarrow J}$ des histogrammes cumulés de la figure 6.11 d'après l'équation 6.12.

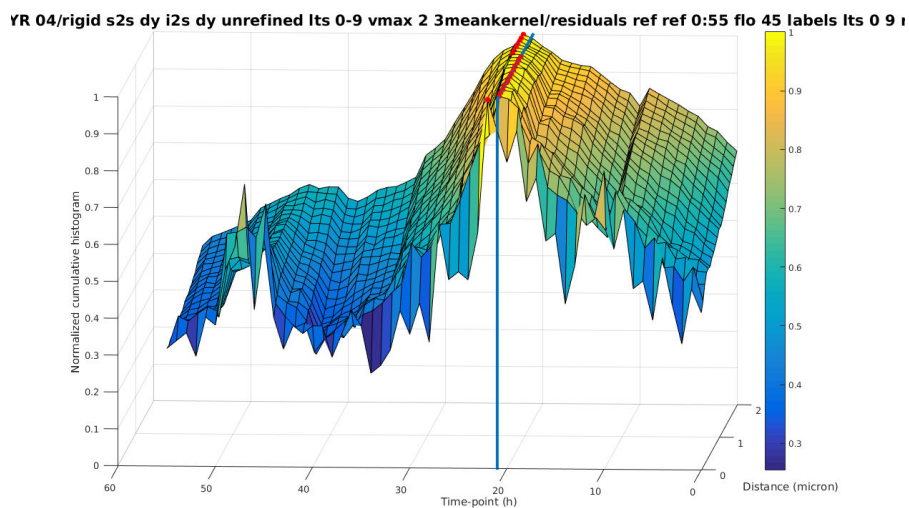


FIGURE 6.13 – Rendu surfacique de l'histogramme cumulé normalisé des distances résiduelles $n_{I(t) \leftarrow J}$ en fonction de $t \in T$ et $d \in D$ (figure 6.12). Les points rouges indiquent pour chaque valeur de $d \in D$ l'instant maximisant la fonction $t \mapsto n_{I(t) \leftarrow J}(d)$ (i.e. vérifiant $n_{I(t) \leftarrow J}(d) = 1$). La ligne bleue correspond à l'instant $t = 21h$.

sur D selon l'équation suivante :

$$s : t \in T \mapsto \frac{1}{\#D} \sum_{d \in D} n_{I(t) \leftarrow J}(d). \quad (6.13)$$

Le domaine d'intégration D des histogrammes cumulés normalisés $n_{I(t) \leftarrow J}$ doit être composé de distances permettant de discriminer les « bons » et les « mauvais » recalages spatiaux. Par exemple, il est inopportun de rechercher l'instant $t \in T$ maximisant l'histogramme cumulé des distances résiduelles à une distance d trop élevée. En effet, le nombre d'appariements dont la distance résiduelle est inférieure à une grande valeur de d ne permettrait alors plus de caractériser de façon pertinente la similarité entre les nuages de points appariés. L'étude de l'épaisseur de la couche L1 des MF montre qu'une cellule de L1 possède une épaisseur caractéristique d'environ $4 \mu\text{m}$ (voir l'annexe D). Nous choisissons donc typiquement de définir D par un ensemble d'une vingtaine de distances uniformément répartis dans un intervalle de distances entre 0 et $2 \mu\text{m}$.

Il est également possible de considérer, à la place de la fonction de similarité définie par l'équation 6.13, la convolution s_v de $t \mapsto s$ avec un noyau de convolution v . Ceci permet le lissage de s (figure 6.14) afin d'assurer une consistance dans l'extraction de l'instant de développement optimal $\hat{t}(J)$:

$$s_v : t \in T \mapsto (s * v)(t). \quad (6.14)$$

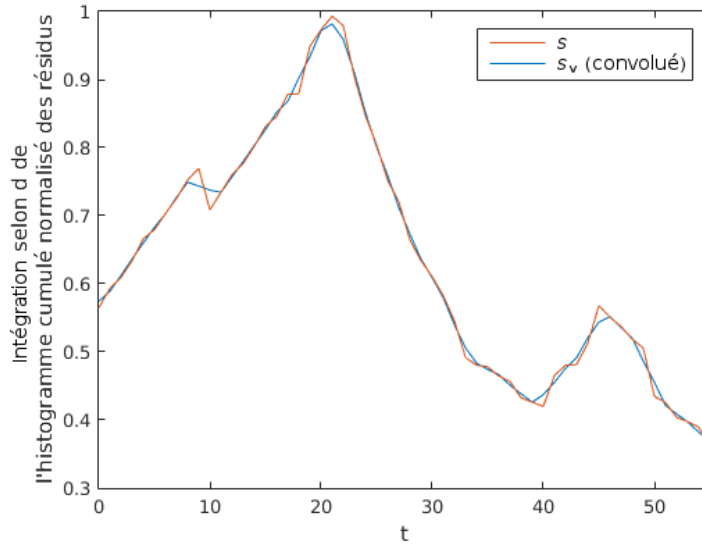


FIGURE 6.14 – Évolution de la fonction de similarité s et de s_v selon $t \in T$. Ici, le noyau de convolution $v = [1, 1, 1]/3$ permet de calculer à chaque instant t la moyenne de s sur trois instants consécutifs. Dans cet exemple, l'instant optimisant l'une ou l'autre des fonctions de similarité vaut $\hat{t}(J) = 21\text{h}$.

6.5 Stratégies de recalage de séquences de MF

Nous avons présenté dans les précédentes sections de ce chapitre une méthode générale de recalage spatial d'images de MF (section 6.3) et de recalage temporel d'une image sur une séquence d'images de MF (section 6.4). Différents choix stratégiques peuvent être adoptés par l'utilisateur pour l'application de ces méthodes au recalage spatio-temporel d'une séquence d'images « flottantes » (terme technique désignant des images mobiles) sur une séquence interpolée de référence (fixe) de MF.

Le recalage spatial entre deux images fait appel à un processus itératif, l'algorithme du plus proche point itéré, qu'on notera dans ce qui suit ICP (*Iterative Closest Points*). L'utilisateur peut :

- choisir les invariants (primitives) à extraire dans les images à recalcer ;
- choisir une transformée initiale pour le procédé d'optimisation ;
- choisir parmi différents schémas d'optimisation du recalage spatial.

La méthode de recalage temporel d'une image flottante sur une séquence (préalablement interpolée) de référence que nous proposons nécessite :

- de recalcer spatialement l'image flottante sur chacune des images de la séquence de référence. Cela implique de considérer une stratégie de « propagation » de l'initialisation de chaque recalage spatial afin d'épargner à l'utilisateur une initialisation manuelle systématique du processus de recalage.
- de choisir le critère de similarité s de l'équation 6.9 à optimiser.

La considération du recalage spatio-temporel des différentes images d'une séquence flottante sur une séquence de référence interpolée nous invite enfin à envisager un schéma de propagation d'une transformée initiale entre les différents instants à recalcer de la série à recalcer.

Dans cette section, nous présentons les différentes stratégies de recalage que nous proposons. La section 6.5.1 introduit les notations que nous utilisons dans la suite de la section. Les stratégies qui concernent le recalage spatial entre deux images sont discutées en section 6.5.2. Puis les stratégies qui concernent le recalage temporel d'une image flottante sur une séquence interpolée de référence sont discutées en section 6.5.3. Enfin, la section 6.5.4 discute de la stratégie de recalage spatio-temporelle d'une séquence d'images flottantes sur une séquence interpolée de référence.

6.5.1 Notations

Dans le but d'alléger les notations utilisées dans les sections précédentes de ce chapitre, on adopte les notations et conventions suivantes pour la suite de la section :

- I_r : image de référence (fixe).
- $I_R = \{I_1, \dots, I_r, \dots, I_n\}$: séquence de référence (interpolée).
- J_f : image flottante (mobile).
- $J_F = \{J_1, \dots, J_f, \dots, J_m\}$: séquence flottante (non interpolée).
- r, f (respectivement R, F) désignent les images I_r, J_f (respectivement les séquences I_R, J_F). Par exemple, $\mathbf{T}_{r \leftarrow f} \stackrel{n}{=} \mathbf{T}_{I_r \leftarrow J_f}$ désigne une transformation de $f \equiv J_f$ vers $r \equiv I_r$.

- \mathbf{T}_x : transformation spatiale.
- μ : méthode de recalage spatial.
- τ : méthode de recalage temporel.
- \tilde{x} : élément de valeur fixée manuellement.
- $x^{(0)}$: initialisation d'une procédure d'optimisation.
- \hat{x} : résultat issu d'une procédure d'optimisation.
- \hat{x}^{meth} : résultat issu d'une procédure d'optimisation soumise à la (ou l'ensemble de) méthode(s) meth.
- $I_r \approx J_f$: superposition de I_r sur J_f satisfaisante au sens d'un contrôle visuel expert.

6.5.2 Recalage image sur image

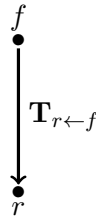


FIGURE 6.15 – Le recalage image sur image consiste à déterminer la transformation spatiale $\mathbf{T}_{r \leftarrow f}$ recalant « au mieux » une image flottante f sur une image de référence r .

6.5.2.1 Méthode de calcul du recalage spatial

D'après les hypothèses de conservation de la forme et du diamètre entre deux MF au même stade de développement, on suppose pouvoir recaler rigide-ment les individus entre eux (section 6.1).

Le recalage spatial consiste à calculer la transformation optimale recalant une image flottante f sur une image de référence r en minimisant une fonction de coût \mathcal{C}_α (section 6.3.2).

Cette partie discute de l'extraction des invariants (primitives) de f et r , et du choix de la paramétrisation de l'algorithme de recalage.

La notation μ désigne une méthode de calcul du recalage spatial intégrant l'extraction des invariants et le procédé d'optimisation.

Extraction des invariants On cherche à exploiter le fait que les interfaces L0/L1, L1/L2 et L2/L3 se conservent au cours du développement des méristèmes. Cela nous amène à discrétiser ces interfaces en nuages de points (section 6.3.1). L'utilisateur peut choisir quelles interfaces il souhaite extraire ainsi que leur niveau de discrétisation (nombre de points). Des opérations peuvent éventuellement être appliquées

sur les interfaces.

En pratique, on extrait toujours l'interface L0/L1 et on peut y ajouter l'information L1/L2. On discrétise chaque interface extraite de l'image flottante f par 5000 points, et de l'image de référence r par de l'ordre de 300000 points.

Lorsqu'on extrait également les interfaces L1/L2, l'hypothèse d'un recalage rigide entre r et f ne vaut que si la distance séparant les interfaces L0/L1 et L1/L2, i.e. l'épaisseur de la L1, est la même pour les deux nuages de points. Pour assurer la validité de cette assertion, on peut appliquer une opération sur l'interface L1/L2 de f visant à ajuster sa distance moyenne à l'interface L0/L1 en nous basant sur les épaisseurs de L1 mesurées sur r et f (annexe D).

Procédé d'optimisation du recalage spatial Le procédé choisi est l'algorithme du plus proche point itéré (*Iterative Closest Points*, ICP). À chaque itération de la méthode, en se basant sur la précédente transformation, on détermine un ensemble de paires de points entre r et f selon le critère du plus proche voisin, puis on calcule la transformation rigide minimisant le résidu quadratique moyen pour ces appariements. On peut choisir d'exclure du calcul de la transformation un pourcentage des plus mauvais appariements (outliers). Cette exclusion renforce la robustesse de l'ICP mais induit le choix d'un paramètre (section 6.3.2).

Aussi, il est possible de choisir de calculer la transformation optimale de f vers r en recherchant :

- la transformation rigide directement (schéma de calcul « direct ») ;
- une translation optimale suivie du calcul de la transformation rigide (schéma de calcul « indirect »).

Par rapport au schéma direct, le schéma de calcul indirect permet d'augmenter progressivement les degrés de liberté de la transformation recherchée, et donc l'espoir de débiter le recalage rigide dans des conditions initiales permettant de converger vers la solution globale du problème de recalage (section 6.5.2.2).

6.5.2.2 Transformation initiale $\mathbf{T}_{r \leftarrow f}^{(0)}$

L'ICP implique de choisir une transformation initiale $\mathbf{T}_{r \leftarrow f}^{(0)}$ qui doit être suffisamment proche de la solution attendue $\hat{\mathbf{T}}_{r \leftarrow f}^\mu$ pour permettre la convergence du processus de recalage μ vers cette solution.

En l'absence d'*a priori*, on propose de déterminer une transformation manuelle $\mathbf{T}_{r \leftarrow f}^{(0)} = \tilde{\mathbf{T}}_{r \leftarrow f}$ telle que $\tilde{\mathbf{T}}_{r \leftarrow f} \circ f \approx r$ à l'aide d'un outil de visualisation synchronisée d'objets 3D tel que Fiji [Schindelin 2012].

6.5.3 Recalage image sur séquence

6.5.3.1 Méthode de calcul du recalage temporel

La notation τ désigne une méthode de calcul du recalage temporel. Actuellement, nous basons le recalage temporel sur l'exploitation de la séquence des distributions

Paramètre	Méthode
μ invariants	L0/L1 $\left(\begin{array}{cc} \text{L0/L1} & \text{L0/L1} \\ \text{L1/L2} & \text{L1/L2 ajusté} \end{array} \begin{array}{cc} \text{L0/L1} & \text{L0/L1} \\ \text{L1/L2} & \text{L1/L2 aj.} \\ \text{L2/L3} & \text{L2/L3 aj.} \end{array} \right)$
μ ICP schéma	Direct Indirect
μ ICP fraction p_α	$p_\alpha \in \left[\frac{1}{2}, 1 \right]$
$\mathbf{T}_{r \leftarrow f}^{(0)}$	Transformation manuelle

TABLE 6.1 – Résumé des différentes stratégies applicables pour le recalage spatial d’une image f sur une image r . La fraction p_α dénote le taux de conservation de points définissant la fonction de coût aux moindres carrés tamisés \mathcal{C}_α (section 5.4.2). Les choix d’invariants placés entre parenthèses sont ceux suggérant l’utilisation d’interfaces en plus de l’interface L0/L1. Bien que nous les mentionnions dans ce tableau, nous n’avons pas pu pleinement exploiter ces différentes possibilités dans nos expériences (se référer à l’annexe D).

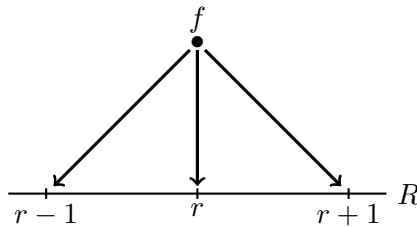


FIGURE 6.16 – Le recalage image sur séquence consiste à déterminer les transformations spatiales $\{\mathbf{T}_{r \leftarrow f}\}_{r \in R}$ recalant spatialement f sur chacune des images de R , puis à déterminer l’instant $\hat{r} \in R$ dont le stade de développement est « le plus proche » de celui de f .

de distances résiduelles du nuage de points flottant vers la référence après le recalage spatial de f sur chaque instant de la séquence de référence R en construisant une fonction de similarité s donnée par l’équation 6.13 (section 6.4.3).

On note \hat{r}^τ le résultat du recalage temporel de f sur R selon τ . Le recalage spatial associé de f vers R est noté $\widehat{\mathbf{T}}_{R \leftarrow f}^\tau = \widehat{\mathbf{T}}_{\hat{r}^\tau \leftarrow f}$.

6.5.3.2 Transformation initiale $\mathbf{T}_{R \leftarrow f}^{(0)}$

Méthode directe On identifie une image $\tilde{r} \in R$ dont le stade de développement est proche de f . On recale f sur \tilde{r} à l'aide de la section 6.5.2, via un procédé de recalage arbitrairement choisi $\tilde{\mu}$. Après contrôle visuel de résultat du recalage, on pose $\mathbf{T}_{R \leftarrow f}^{(0)} = \widehat{\mathbf{T}}_{\tilde{r} \leftarrow f}^{\tilde{\mu}}$.

Méthode affinée On peut affiner la méthode directe en réalisant un recalage préliminaire de f sur R qui utilise l'initialisation $\widehat{\mathbf{T}}_{\tilde{r} \leftarrow f}^{\tilde{\mu}}$ suivie d'une méthode de recalage μ_a pouvant par exemple être un recalage de type translation. Le recalage de f sur R via μ_a et τ résulte en les optima $\widehat{\mathbf{T}}_{R \leftarrow f}^{\{\mu_a, \tau\}}$ et $\hat{r}^{\{\mu_a, \tau\}}$. On pose ainsi $\mathbf{T}_{R \leftarrow f}^{(0)} = \widehat{\mathbf{T}}_{R \leftarrow f}^{\{\mu_a, \tau\}}$.

6.5.3.3 Schéma de propagation de la transformation initiale : $\mathbf{T}_{r \leftarrow f}^{(0)}$

Il s'agit de déterminer $\mathbf{T}_{r \leftarrow f}^{(0)}$ pour tout $r \in R$ sachant $\mathbf{T}_{R \leftarrow f}^{(0)}$ et éventuellement un point de départ $r^{(0)}$ pour le schéma.

Schéma statique

$$\mathbf{T}_{r \leftarrow f}^{(0)} = \mathbf{T}_{R \leftarrow f}^{(0)}, \quad \forall r \in R. \quad (6.15)$$

Pour chaque image de R , on utilise la même initialisation $\mathbf{T}_{r \leftarrow f}^{(0)}$. Cela peut se justifier si on considère que les images de la séquence R sont auparavant recalées rigidement entre elles.

Schéma dynamique par propagation

$$\mathbf{T}_{r \leftarrow f}^{(0)} = \begin{cases} \mathbf{T}_{R \leftarrow f}^{(0)}, & \text{si } r = r^{(0)}, \\ \widehat{\mathbf{T}}_{r+1 \leftarrow f}, & \text{si } r < r^{(0)}, \\ \widehat{\mathbf{T}}_{r-1 \leftarrow f}, & \text{si } r > r^{(0)}. \end{cases} \quad (6.16)$$

Le choix de l'image initiale $r^{(0)}$ dépend du choix de la méthode d'initialisation de la transformation. Pour la méthode directe, on pose $r^{(0)} = \tilde{r}$. Pour la méthode affinée, on pose $r^{(0)} = \hat{r}^{\{\mu_a, \tau\}}$.

Ce schéma permet de mettre à jour la transformation initiale de proche en proche en partant du temps initial $r^{(0)}$. Ce schéma d'initialisation doit permettre une plus grande adaptabilité du recalage aux déplacements et variations de morphologie de R qu'avec le schéma statique. L'inconvénient majeur de cette méthode est que dès qu'un recalage spatial de la série est défectueux, l'initialisation du recalage pour toutes les images suivantes de la série sera *a priori* mauvaise et conduira l'algorithme de recalage itératif vers un optimum local différent de celui recherché. Si cet échec survient avant le recalage de l'image optimale escomptée, la possibilité de recalculer spatialement f sur cette image de manière correcte devient alors très incertaine.

Paramètre	Méthode	
τ $D = [d_{\min}, d_{\max}]$	$d_{\min} = 0 \mu\text{m}$ $d_{\max} \in [1 \mu\text{m}, 4 \mu\text{m}]$	
τ critère de similarité	Choix d'un masque de convolution de s	
$\mathbf{T}_{R \leftarrow f}^{(0)}$	Direct	Affiné ($\mu_{\text{translation}}$)
Schéma de propagation	Statique	Dynamique

TABLE 6.2 – Résumé des différentes stratégies applicables pour le recalage temporel d'une image f sur une séquence interpolée d'images de référence R .

6.5.4 Recalage séquence sur séquence

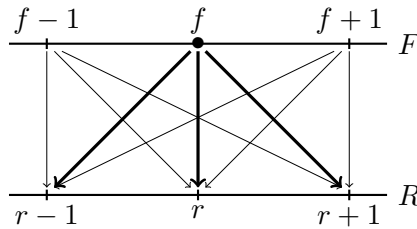


FIGURE 6.17 – Le recalage séquence sur séquence consiste à déterminer le recalage spatio-temporel de chacune des images de la séquence flottante F sur la séquence de référence R .

6.5.4.1 Transformation initiale $\mathbf{T}_{R \leftarrow F}^{(0)}$

On identifie une image $\tilde{f} \in F$ que l'on recalc sur R à l'aide de la section 6.5.3, via des procédés de recalage spatio-temporels $\tilde{\mu}$ et $\tilde{\tau}$ choisis afin d'obtenir $\hat{\mathbf{T}}_{R \leftarrow \tilde{f}}^{\{\tilde{\mu}, \tilde{\tau}\}} \circ \tilde{f} \approx \hat{r}^{\tilde{\tau}}$. On pose alors $\mathbf{T}_{R \leftarrow F}^{(0)} = \hat{\mathbf{T}}_{R \leftarrow \tilde{f}}^{\{\tilde{\mu}, \tilde{\tau}\}}$.

6.5.4.2 Schéma de propagation de la transformation initiale : $\mathbf{T}_{R \leftarrow f}^{(0)}$

Il s'agit de déterminer $\mathbf{T}_{R \leftarrow f}^{(0)}$ pour tout $f \in F$ sachant $\mathbf{T}_{R \leftarrow F}^{(0)}$ et un éventuel point de départ $f^{(0)}$.

Schéma statique

$$\mathbf{T}_{R \leftarrow f}^{(0)} = \mathbf{T}_{R \leftarrow F}^{(0)}, \quad \forall f \in F. \quad (6.17)$$

Paramètre	Méthode	
Schéma de propagation	Statique	Dynamique

TABLE 6.3 – Résumé des différentes stratégies applicables pour le recalage spatio-temporel de chacune des images d’une séquence flottante F sur une séquence interpolée de référence R .

Ce schéma présuppose que les images de F ainsi que de R soient rigidement recalées 2 à 2.

Schéma dynamique par propagation

$$\mathbf{T}_{R \leftarrow f}^{(0)} = \begin{cases} \mathbf{T}_{R \leftarrow F}^{(0)}, & \text{si } f = f^{(0)}, \\ \widehat{\mathbf{T}}_{R \leftarrow f+1}, & \text{si } f < f^{(0)}, \\ \widehat{\mathbf{T}}_{R \leftarrow f-1}, & \text{si } f > f^{(0)}. \end{cases} \quad (6.18)$$

On choisit ici $f^{(0)} = \tilde{f}$ par défaut.

De même que pour le recalage d’une image sur une séquence, la limitation du schéma dynamique est qu’en cas de mauvais résultat de recalage pour une des images de la séquence F , cela entraînera de manière presque certaine l’échec du recalage des images suivantes de F sur R .

Le schéma d’initialisation dynamique suppose que les changements morphologiques entre deux images consécutives f et $f + 1$ sont suffisamment faibles pour que la transformation recalant l’une sur R soit une bonne initialisation pour l’autre au sens de la section 6.5.2.2.

Une nuance apparaît toutefois entre le schéma dynamique par propagation du recalage séquence sur séquence et celui du recalage image sur séquence. En effet, contrairement à la séquence R , la séquence flottante F n’est pas interpolée, ce qui signifie que le délai important séparant les acquisitions de deux images consécutives f et $f + 1$ de F peut potentiellement conduire à des changements morphologiques conséquents.

6.6 Expériences et résultats

6.6.1 Données

6.6.1.1 Acquisition des données

Nous avons pris des plants d’*Arabidopsis thaliana* sauvages, marqués génétiquement par une protéine fluorescente jaune (YFP) acylée, pour la génération des données imagées. La YFP acylée marque clairement les murs cellulaires en vue d’une segmentation des cellules. Les plants se développent initialement dans de la terre en

bénéficiant d'un éclairage continu, jusqu'à ce qu'ils basculent de la phase végétative à la phase reproductrice de leur développement.

Afin de pouvoir être observées au microscope, les plantules dont la tige d'inflorescence ne dépasse pas une taille de 1 cm sont transférées dans une boîte en plastique leur offrant un milieu de Murashige et Skoog (MS) contenant tous les nutriments nécessaires au développement végétal. Le transfert est réalisé en prêtant attention à préserver les racines des plantules. Une dissection du méristème permet d'éliminer les éléments végétaux obstruant la visibilité des MF.

Les plants sont finalement imagés à l'aide d'un microscope confocal vertical (Zeiss LSM 780 ou LSM 700) équipé d'un objectif 20x à immersion dans l'eau. Les acquisitions sont réalisées au niveau d'une région siège de la formation d'un nouveau méristème floral. Les données sont acquises à différents stades du développement des MF observés, constituant pour chaque MF une séquence temporelle d'images 3D avec de 6 à 18 images par séquence. Les acquisitions sont séparées de 4 à 13 heures. Entre deux acquisitions, les plants sont maintenues éclairés.

Lors de chaque acquisition d'image, une pile d'images 2D (vues en coupe de l'objet observé) est capturée selon l'axe vertical z du microscope, permettant de constituer un volume 3D. Chaque image 3D est construite à partir de 130 à 683 acquisitions d'images 2D. La résolution verticale de l'acquisition peut varier de 130 à 610 nm. Une coupe 2D est typiquement une image de taille 400×400 avec une résolution de pixel d'environ 250 nm selon les axes x et y du microscope.

En raison du déplacement vertical du bouton floral dans le champ de vue au cours de l'acquisition des images, dû à la croissance de la plante, une estimation de la résolution en z du microscope est nécessaire. Pour cela, un double balayage rapide du spécimen est réalisé, l'un rapide et l'autre plus lent. La comparaison des deux images acquises permet d'estimer la résolution selon z de l'image obtenue par balayage lent.

Pour pouvoir extraire les primitives de chaque image (section 6.3.1), les segmentations de cellules de chaque séquence 3D+t sont obtenues à l'aide de l'algorithme MARS [Fernandez 2010a] puis corrigées manuellement si besoin.

6.6.1.2 Détail des données

Nous avons travaillé avec quatre séquences temporelles d'images 3D de méristèmes floraux, nommément :

- YR_01 : 18 instants imagés toutes les 4 à 8 h. Durée totale d'acquisition : 132 heures.
- PlantC_P1 : 6 instants imagés toutes les 10 à 13 h. Durée totale d'acquisition : 55 heures.
- PlantE_P1 : 7 instants imagés toutes les 5 à 13 h. Durée totale d'acquisition : 55 heures.
- PlantE_P2 : 7 instants imagés toutes les 5 à 13 h. Durée totale d'acquisition : 55 heures.

Notons que les MF PlantE_P1 et PlantE_P2 appartiennent à la même plantule. La figure 6.18 présente une vue du dessus de chaque image des différentes séquences temporelles.

6.6.2 Interpolation d'images

6.6.2.1 Étape de recalage

L'étape d'affinage de la résolution temporelle d'une séquence d'images de MF, présentée en section 6.2, consiste à construire une séquence d'images interpolées à partir des images d'origine. Cette étape repose principalement sur la capacité de recalibrer spatialement deux images consécutives d'une séquence originale de MF. Nous proposons d'appliquer un recalage spatial en trois temps (figure 6.19) :

- d'abord, on compense les mouvements rigides de la plante dans le champ de vue à l'aide d'un recalage rigide ;
- puis on réalise un recalage linéaire (affine) entre les deux images rigidement recalées ;
- enfin, on recalc non-linéairement les images en calculant un champ de déformation (champ de vecteurs) permettant de superposer les murs cellulaires des deux images.

Ces trois étapes de calcul sont réalisées à l'aide de la méthode de recalage par blocs [Ourselin 2000].

L'algorithme de recalage par blocs, qui consiste à construire itérativement et selon plusieurs niveaux hiérarchiques (système dit « pyramidal ») des correspondances entre des blocs d'images (primitives « iconiques »), dépend d'un nombre relativement important de paramètres que l'utilisateur doit fixer. Les paramètres permettent de régir en particulier les niveaux de pyramides, la construction des primitives iconiques, les hypothèses d'appariement des primitives ou encore des critères d'élasticité imposés au champ de déformation recherché (pour le recalage non-linéaire seulement). Le lecteur intéressé peut se référer à une documentation quelque peu étayée relative à l'utilisation de cet outil de recalage [Malandain 2016].

L'expérience nous montre que les recalages rigide et affine appliqués à deux images consécutives d'une séquence d'images de MF sont satisfaisants au sens d'une évaluation visuelle sur l'ensemble de la base de données traitée, pour un total de 34 paires d'images testées, avec une faible sensibilité au choix des paramètres de recalage.

Le recalage non-linéaire par blocs est davantage sensible à sa paramétrisation. Dans nos expériences, nous avons choisi un jeu de paramètres « standard » qui nous a permis de recalibrer 28 des 34 paires d'images testées en satisfaisant un contrôle visuel du résultat. L'étude des 6 paires d'images qui n'ont pas satisfait ce contrôle montre de manière attendue que les difficultés de recalage surviennent lorsque les transformations morphologiques subies par un MF entre deux acquisitions sont fortes, en particulier lorsque le stade de formation des sépales est atteint par le MF. Nous sommes finalement parvenus à recalibrer non-linéairement les 6 paires d'images

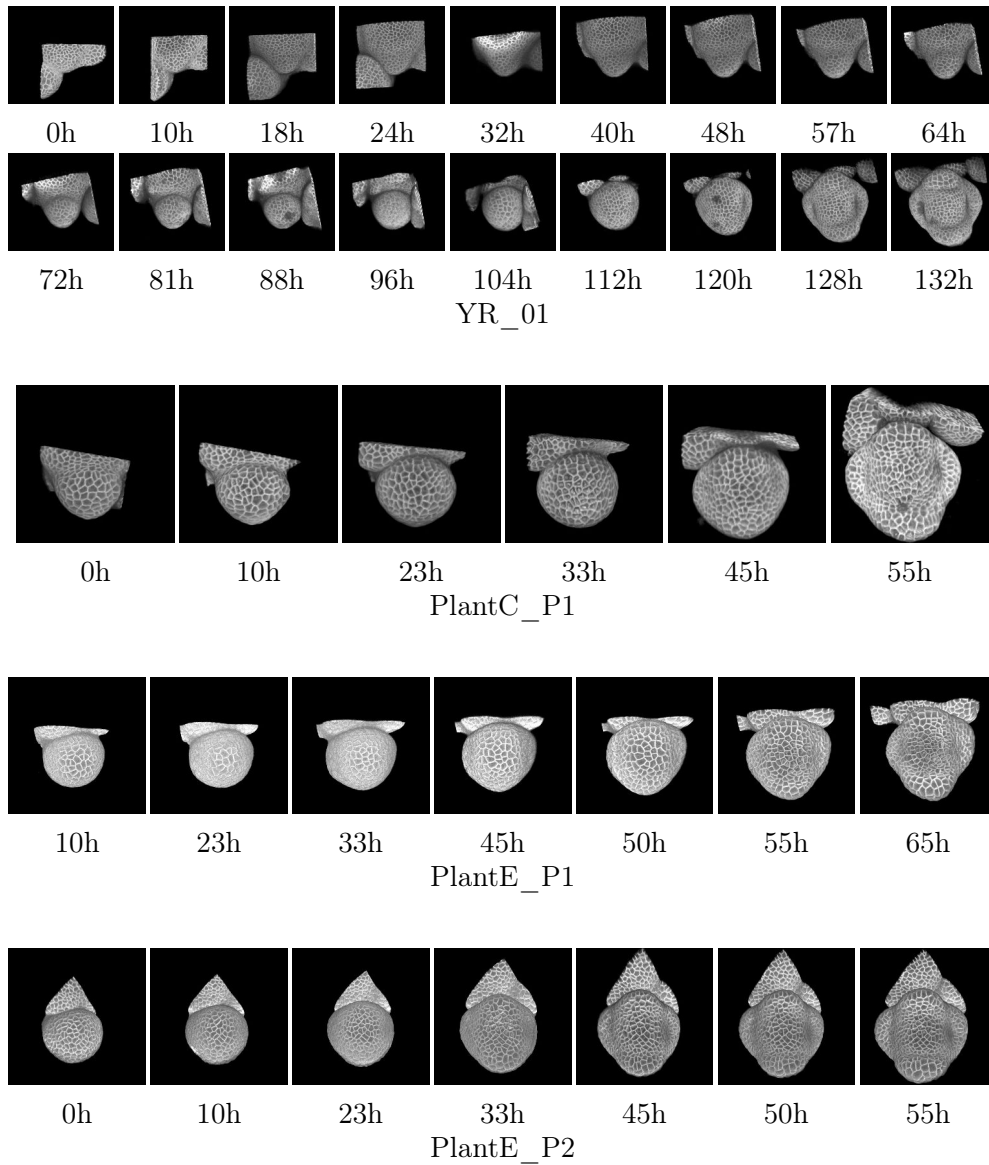


FIGURE 6.18 – Rendus 3D vus de dessus des volumes d’images capturés des MF YR_01, PlantC_P1, PlantE_P1 et PlantE_P2. Ici, les déplacements rigides des méristèmes dans le champ de vue sont compensés (section 6.2.1) et les contrastes ajustés sur une même séquence d’images (section 6.2.2).

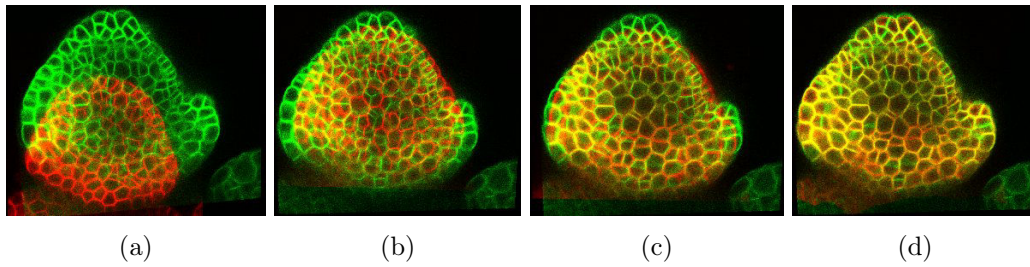


FIGURE 6.19 – Étapes de recalage entre deux images consécutives d’une séquence temporelle d’images de MF. Le canal vert (respectivement rouge) présente une vue en coupe de l’image de la séquence YR_01 à l’instant $t_{i+1} = 128\text{h}$ (respectivement à l’instant $t_i = 120\text{h}$) au cours des différentes étapes de recalage. (a) Images originales. (b) Images recalées rigidement. (c) Images recalées linéairement (i.e. recalage affine). (d) Images recalées non-linéairement.

en considérant des jeux de paramètres spécifiques pour la méthode de recalage par blocs.

6.6.2.2 Interpolation d’images segmentées

La méthode d’interpolation d’images segmentées introduite en section 6.2.3.3 et détaillée en annexe C a été appliquée sur les quatre séquences d’images segmentées.

Nous présentons en figure 6.20 un exemple de résultat d’interpolation de deux images segmentées selon cette méthode illustrant la capacité de la méthode d’interpolation à construire des segmentations compatibles entre elles. On remarque en effet sur cette figure que les interpolations de cellules segmentées des deux images d’origine appartenant à une même lignée cellulaire se superposent parfaitement au niveau des frontières de leurs cellules.

La figure 6.21 illustre le résultat de l’interpolation de segmentations réalisée sur une séquence entière. On observe dans cette figure la seconde propriété de la méthode d’interpolation proposée, à savoir la continuité du résultat de l’interpolation sur l’ensemble de la séquence segmentée.

6.6.3 Recalage spatial d’images de MF

6.6.3.1 Sélection des primitives pour le recalage spatial

Nous avons présenté en section 6.3.1 notre méthode d’extraction de primitives géométriques pour le recalage spatial de deux images de MF I et J . Les primitives que nous utilisons sont des nuages de points répartis aux interfaces entre les différentes couches de cellules de chaque MF. Les nuages de points extraits de I et J sont respectivement notés P_k^I et P_k^J où $k \in \{1, 2, 3\}$ désigne l’indice correspondant à l’interface L0/L1 ($k = 1$), L1/L2 ($k = 2$) ou L2/L3 ($k = 3$). L’algorithme des plus proches points itérés permet de calculer la transformation spatiale minimisant

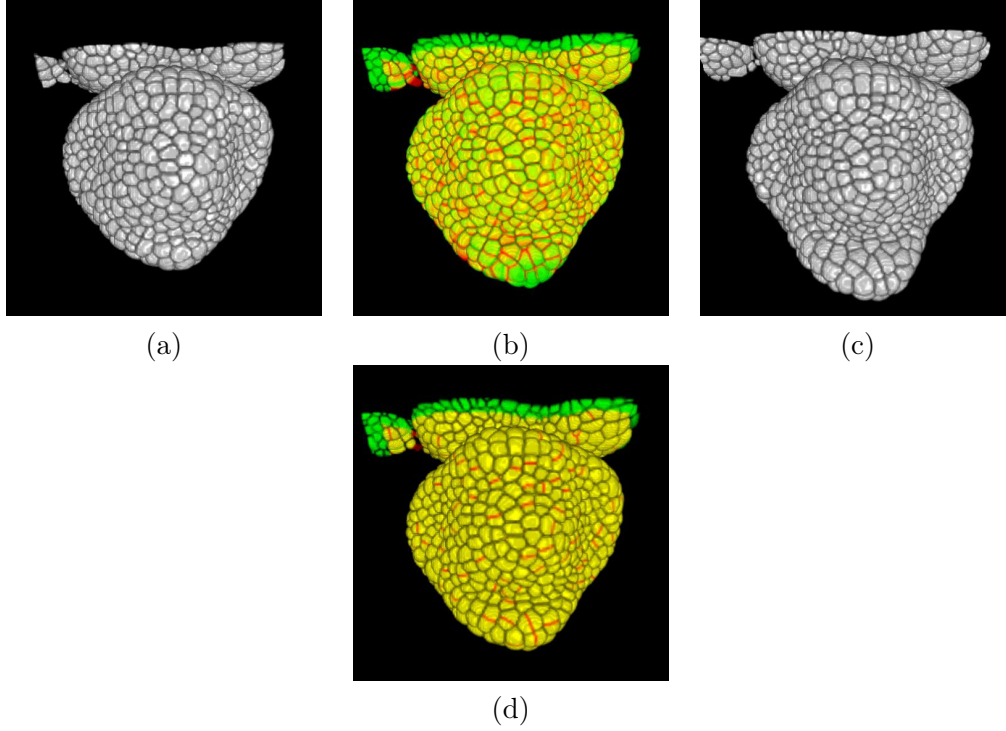


FIGURE 6.20 – Exemple d’interpolation de segmentations réalisée sur une paire d’images segmentées de la séquence PlantE_P1. (a) Image segmentée à l’instant $t_i = 55$ h. (c) Image segmentée à l’instant $t_{i+1} = 65$ h. (b) Superposition de l’image segmentée (a) transformée et l’image segmentée (c) transformée (respectivement en rouge et vert) à l’instant intermédiaire $t = \frac{t_i+t_{i+1}}{2}$ calculées à partir des champs de transformation intermédiaires $\mathbf{T}_{t_i \leftarrow t}$ et $\mathbf{T}_{t_{i+1} \leftarrow t}$ (section 6.2.3.1). (d) Superposition des segmentations interpolées de (a) et (b) (respectivement en rouge et vert) à l’instant t (section 6.2.3.3). On observe en (b) que les segmentations directement transformées ne sont pas cohérentes entre elles au sens où les frontières de cellules d’une même lignée ne coïncident pas parfaitement entre elles. En (d), on observe que les segmentations qui résultent de la méthode d’interpolation proposée sont rendues cohérentes dans tout le domaine de l’image défini par l’intersection entre les champs de vue respectifs des deux images segmentées. Les divisions cellulaires survenues entre les instants t_i et t_{i+1} apparaissent en rouge dans (d).

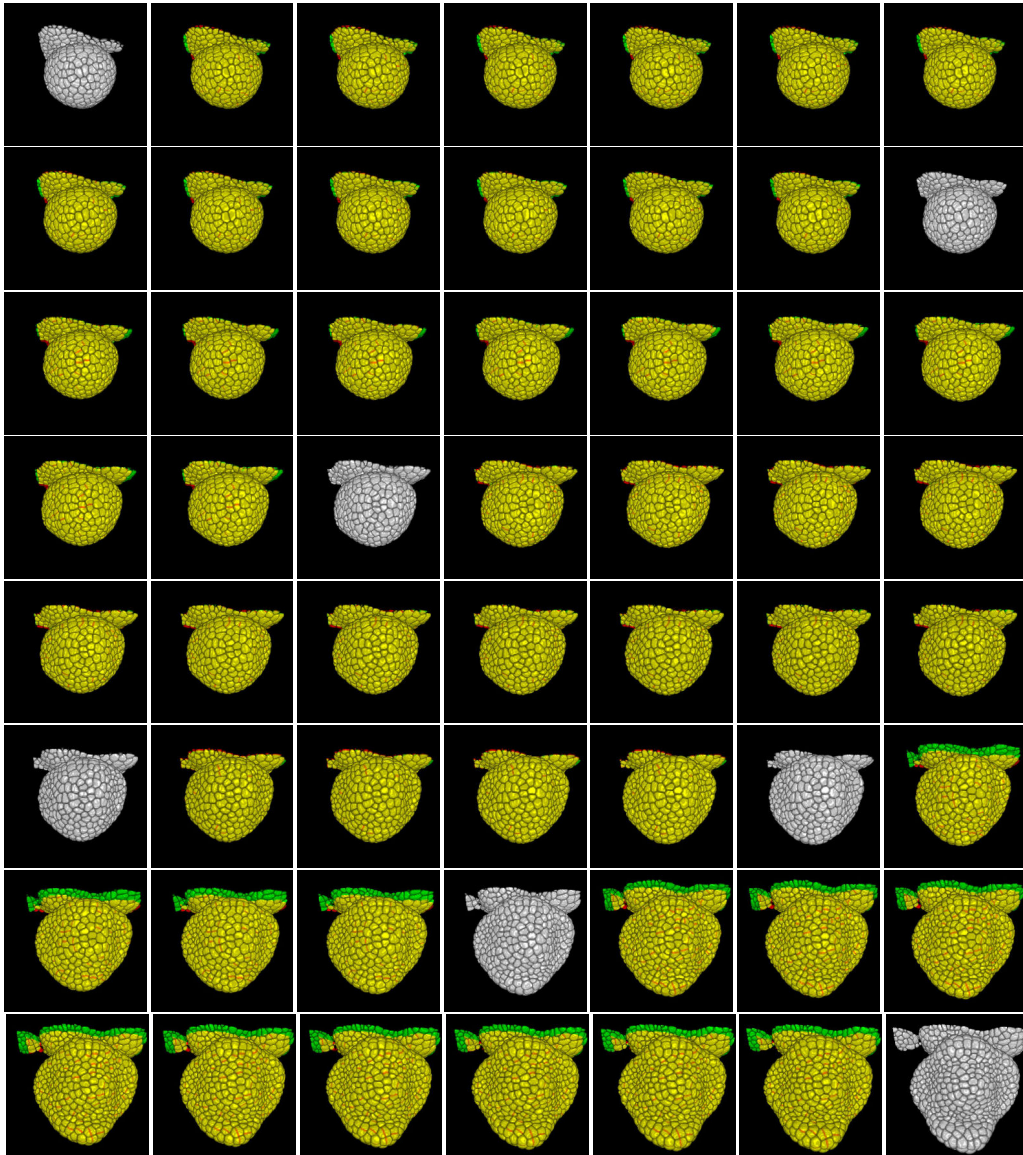


FIGURE 6.21 – Résultat de la construction d’une séquence interpolée d’images segmentées avec un pas de 1 h entre chaque image à partir de la séquence d’origine PlantE_P1. Les images segmentées des instants d’origine de la séquence (10h, 23h, 33h, 45h, 50h, 55h, 65h) apparaissent en blanc. Pour chaque instant intermédiaire, nous affichons la superposition des segmentations interpolées provenant des instants antérieur (en rouge) et postérieur (en vert) à cet instant de la séquence d’origine.

la distance résiduelle moyenne entre les points de $P_K^I = \{P_k^I\}_{k \in K}$ et $P_K^J = \{P_k^J\}_{k \in K}$ appartenant à une même interface d'indice $k \in K$.

L'hypothèse de conservation de la forme globale et de la taille des méristèmes floraux à un même instant de développement nous a amenés à considérer un recalage de type rigide, c'est-à-dire sans déformation. Cependant, cette hypothèse n'est a priori valide qu'à la surface des MF, c'est-à-dire à l'interface L0/L1. À l'intérieur des méristèmes, les hypothèses d'invariance peuvent être discutées. Il est reconnu que les couches de cellules L1, L2 et L3 respectent une concentricité, ce qui permet de considérer que l'hypothèse de conservation de la forme globale est également valide pour les interfaces L1/L2 et L2/L3 de MF. En revanche, l'état de l'art ne permet pas d'affirmer que les couches de cellules possèdent une même épaisseur selon les individus. À partir des données que nous possédons, nous avons réalisé des mesures d'épaisseur locale des cellules de couche L1 sur les images de chaque séquence de MF. Cette étude, présentée en annexe D, établit de façon évidente que l'épaisseur « moyenne » de la couche L1 varie de façon marquée au cours du développement d'un même MF ainsi que d'un individu à l'autre. Il n'est donc a priori pas justifié de chercher à recalibrer rigide les nuages de points P_k^I et P_k^J entre eux pour $k \geq 2$.

Ainsi, le choix de l'ensemble $K \subset \{1, 2, 3\}$ des indices d'interfaces à utiliser pour le recalage rigide de I et J n'est pas trivial, bien que les problèmes de recalage tirent le plus souvent profit de l'exploitation de primitives de nature à contraindre plus fortement le problème.

Nous avons étudié l'influence du paramètre K sur un couple d'images provenant des MF YR_01 à $t = 120h$ et PlantE_P1 à $t = 48h$ identifiés comme étant à un même stade de développement en recalant rigide ces images à partir d'une initialisation $\mathbf{T}_{I \leftarrow J}^{(0)}$ et avec un paramètre de tamisage $p_\alpha = 0,9$ communs. Le contrôle visuel attentif des recalages spatiaux obtenus pour $K = \{1, 2\}$ et $K = \{1\}$ montre que la considération de l'interface L1/L2 dégrade qualitativement le résultat du recalage pour ce couple d'images (voir la figure 6.22).

6.6.3.2 Influence de l'occultation du MF par le champ de vue

L'une des difficultés liées au problème d'extraction des primitives des objets que l'on souhaite recalibrer est que dans le cas des plantes, les organismes ne sont jamais entièrement imagés. Le champ de vue du microscope « masque » à sa base le méristème floral à une hauteur variable. La région de la plante située au point de rattachement du MF avec le méristème apical caulinaire (MAC) est elle aussi partiellement masquée. On parlera d'occultation partielle du MF par le champ de vue. La figure 6.23 montre l'impact négatif d'une forte occultation d'une segmentation de MF sur la qualité du recalage spatial du MF YR_01 à l'instant $t = 120h$ sur le MF PlantE_P1 à $t = 48h$. On met ainsi en évidence qu'il est préférable de recalibrer des objets faiblement occultés, ce qui peut s'avérer contraignant lors de l'étape d'acquisition des séquences temporelles d'images de MF.

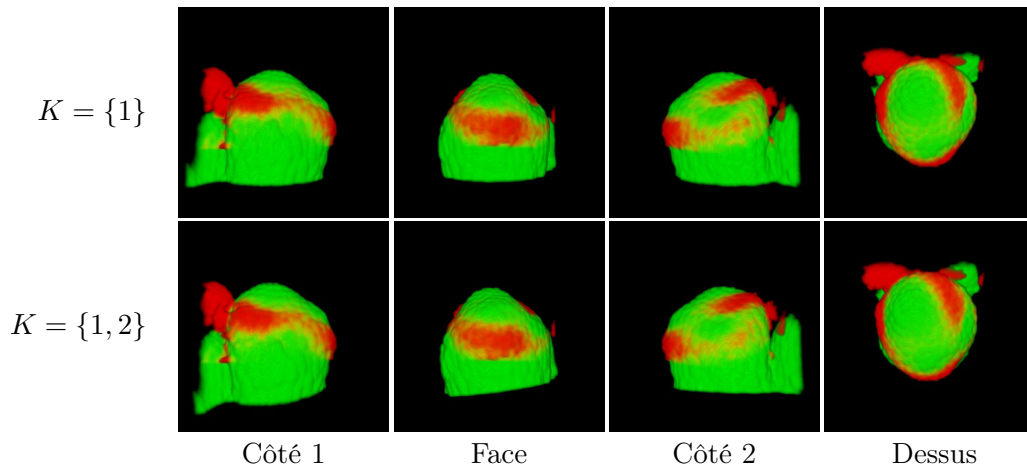


FIGURE 6.22 – Vues 3D du résultat du recalage de l’instant $t = 120h$ du MF YR_01 (rouge) sur l’instant $t = 48h$ du MF PlantE_P1 (vert) pour différentes extractions de primitives (paramétrées par K).

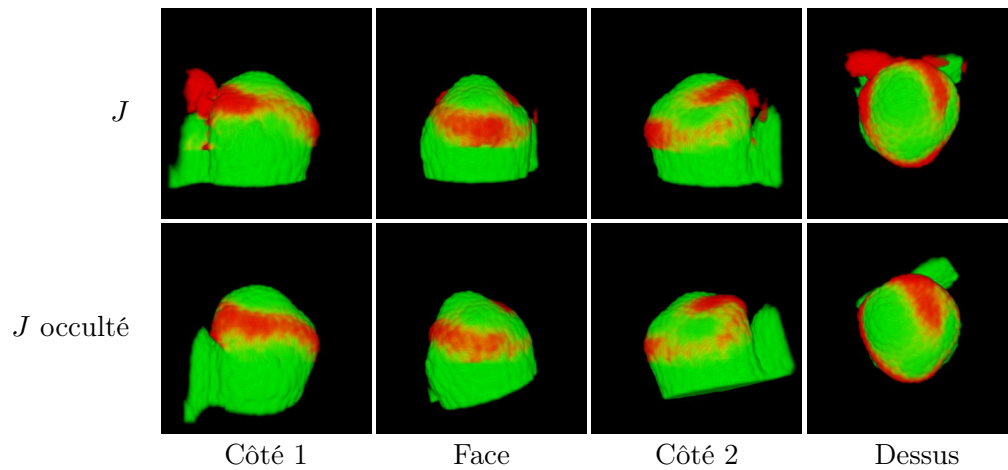


FIGURE 6.23 – Vues 3D du résultat du recalage de J correspondant à l’instant $t = 120h$ du MF YR_01 (rouge) sur I correspondant à l’instant $t = 48h$ du MF PlantE_P1 (vert) selon deux niveaux d’occlusion du MF segmenté J . En haut : segmentation d’origine de J . En bas : occultation de la région correspondant à la partie du méristème apical caulinaire (MAC) en contact avec le MF J . Recalages calculés avec $K = \{1\}$ et $p_\alpha = 0,9$.

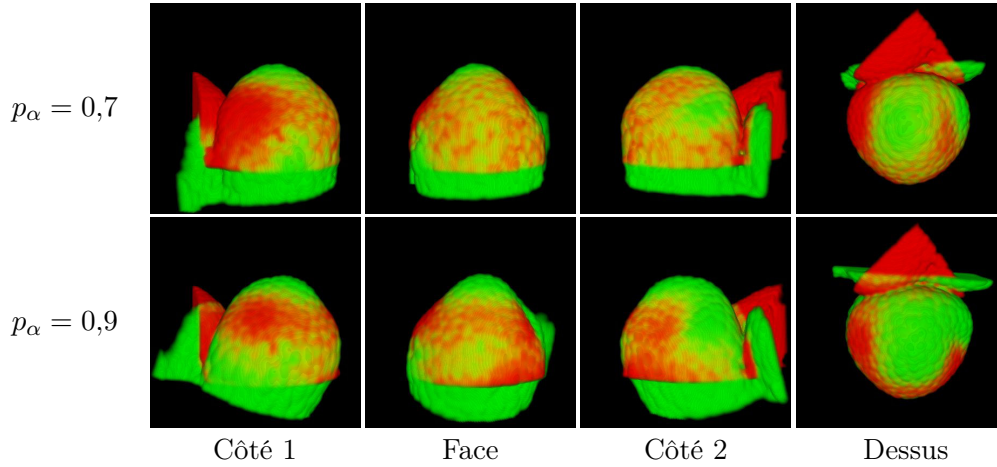


FIGURE 6.24 – Vues 3D du résultat du recalage de l’instant $t = 10\text{h}$ du MF PlantE_P2 (rouge) sur l’instant $t = 39\text{h}$ du MF PlantE_P1 (vert) pour différents niveaux de tamisage. En haut : $p_\alpha = 0,7$. En bas : $p_\alpha = 0,9$. Recalages calculés avec $K = \{1\}$.

6.6.3.3 Influence du coefficient de tamisage

On montre dans la figure 6.24 que le paramètre de tamisage p_α (section 6.3.2) peut avoir une influence sur le résultat du recalage spatial de deux images. Dans cet exemple, les données utilisées sont le MF PlantE_P2 à $t = 10\text{h}$ (J) et le MF PlantE_P1 à $t = 39\text{h}$ (I). Le recalage du nuage de points P_K^J sur P_K^I est calculé pour $K = \{1\}$ (recalage des interfaces L0/L1 seulement). La comparaison des résultats du recalage spatial de J sur I pour un paramètre p_α valant 0,7 et 0,9 montre qualitativement un meilleur recalage pour $p_\alpha = 0,7$, ce qui montre l’importance de considérer un algorithme de recalage par plus proches points itérés robuste. Cependant, des expériences de recalage d’images de la séquence PlantE_P1 sur la séquence PlantE_P2 montrent à l’inverse un meilleur comportement de l’algorithme de recalage spatial pour une valeur de p_α de l’ordre de 0,9 (voir section 6.6.4).

Afin d’illustrer l’effet du paramètre p_α sur l’optimisation du recalage, nous avons représenté dans la figure 6.25 les distances résiduelles associées aux points de P_K^J recalés sur P_K^I pour les deux images tests de la figure 6.24, ainsi que la répartition des points de P_K^J rejetés par l’algorithme de recalage (*outliers*) lors du calcul de la transformée optimale $\hat{\mathbf{T}}_{I \leftarrow J}$. On voit à l’aide de cette figure que le phénomène de rotation indésirable de J par rapport à I pour $p_\alpha = 0,9$ est dû à la prise en considération de points de P_K^J provenant d’une région qui n’appartient pas au MF mais au MAC et qui est davantage occultée dans la segmentation de référence I . Le choix de $p_\alpha = 0,7$ permet de ne plus considérer cette région lors du calcul de la fonction de coût \mathcal{C}_α (équation 5.15). On observe alors qualitativement un recalage spatial de bien meilleure qualité.

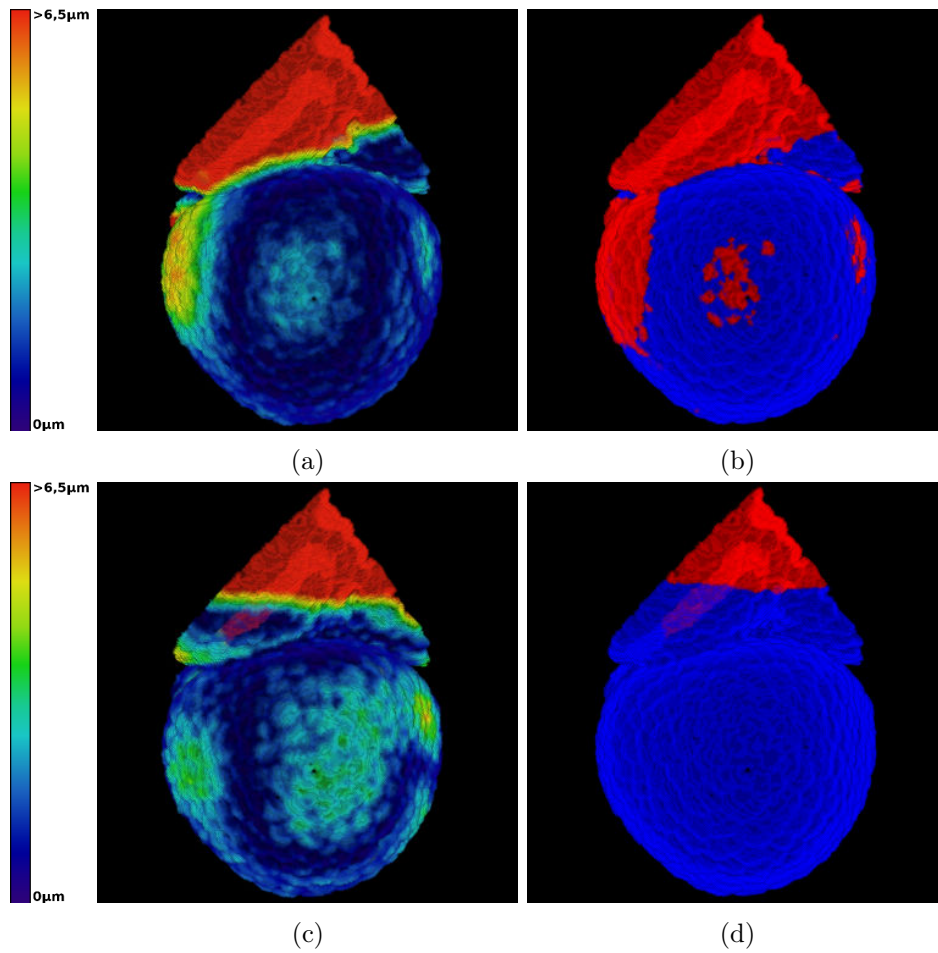


FIGURE 6.25 – Distribution des distances résiduelles et des points de P_K^J rejetés pour le calcul de la transformée optimale de J vers I dont le résultat est illustré en figure 6.24. Calculs réalisés selon $p_\alpha = 0,7$ pour (a,b) et $p_\alpha = 0,9$ pour (c,d). Le paramètre $K = \{1\}$ est fixé.

6.6.4 Stratégies de recalage

La section 6.5 présente la déclinaison de la méthode de recalage spatio-temporel de séquences d'images de méristèmes floraux selon plusieurs choix de paramétrisation et « stratégies » de propagation de l'initialisation du recalage spatial d'image en image. Ce dernier point est en effet crucial dans la mesure où une initialisation correcte est indispensable pour permettre à la méthode itérative de recalage de converger vers une solution jugée satisfaisante, et qu'une initialisation manuelle de la méthode pour chaque paire d'images à recaler considérée constituerait une tâche extrêmement fastidieuse lorsqu'on considère les problèmes de recalage image sur séquence et séquence sur séquence.

Notons \mathcal{M} un ensemble de stratégies de recalage. Soit $M \in \mathcal{M}$ une stratégie de recalage donnée, F une séquence d'images flottantes et R une séquence interpolée d'images de référence. Pour toute image $f \in F$, on note respectivement \hat{r}_f^M et $\hat{\mathbf{T}}_{R \leftarrow f}^M$ l'image extraite de R et la transformation spatiale recalant f sur cette image issues de la méthode de recalage à l'aide de la stratégie M . Nous avons réalisé des expériences de recalage séquence sur séquence selon un ensemble de stratégies \mathcal{M} , et ainsi construit pour chaque paire de séquences considérée des familles de recalages $\{(\hat{r}_f^M, \hat{\mathbf{T}}_{R \leftarrow f}^M)\}_{f \in F, M \in \mathcal{M}}$ à partir d'une initialisation commune.

Pour évaluer chacun de ces recalages, nous les comparons à des recalages manuellement estimés que l'on note $\{(\tilde{r}_f, \tilde{\mathbf{T}}_{R \leftarrow f})\}_{f \in F}$.

Nous effectuons la comparaison d'un recalage $(\hat{r}, \hat{\mathbf{T}})$ avec le recalage expertisé $(\tilde{r}, \tilde{\mathbf{T}})$ en considérant la mesure de :

- $\Delta t(f) = |t(\hat{r}) - t(\tilde{r})|$, où $t(r)$ est l'instant de développement de $r \in R$;
- $\hat{\mathbf{T}} \circ (\tilde{\mathbf{T}}^{-1}) = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{pmatrix}$, où \mathbf{R} et \mathbf{t} représentent respectivement la rotation et la translation résiduelles de $\hat{\mathbf{T}}$ par rapport à $\tilde{\mathbf{T}}$, et desquels on peut mesurer :
 - $\Delta\theta(f) = \arccos\left(\frac{\text{trace}(\mathbf{R})-1}{2}\right)$, représentant l'écart angulaire entre les rotations de $\hat{\mathbf{T}}$ et $\tilde{\mathbf{T}}$;
 - $\Delta\text{trans}(f) = \|\mathbf{t}\|$, norme de la translation résiduelle entre $\hat{\mathbf{T}}$ et $\tilde{\mathbf{T}}$.

Pour une paire donnée de séquences F et R et pour différents choix de $M \in \mathcal{M}$, nous effectuons les mesures de comparaison entre les familles de recalages $\{(\hat{r}_f^M, \hat{\mathbf{T}}_{R \leftarrow f}^M)\}_{f \in F}$ et les recalages expertisés $\{(\tilde{r}_f, \tilde{\mathbf{T}}_{R \leftarrow f})\}_{f \in F}$. Pour une stratégie M donnée, un « bon » recalage spatio-temporel d'une image flottante f sur une séquence R suppose que la comparaison de $(\hat{r}_f^M, \hat{\mathbf{T}}_{R \leftarrow f}^M)$ avec $(\tilde{r}_f, \tilde{\mathbf{T}}_{R \leftarrow f})$ conduit à des mesures de $\Delta t(f)$ et $\Delta\theta(f)$ de faible ampleur, et donc à un couple $(\Delta t(f), \Delta\theta(f))$ voisin de $(0, 0)$. Le recalage d'une séquence F sur une séquence R sera ainsi d'autant « meilleur » que la distribution de points $\{(\Delta t(f), \Delta\theta(f))\}_{f \in F}$ sera regroupée vers le point de coordonnées $(0, 0)$. Chaque graphe de la figure 6.26 affiche pour un paramètre de tamisage p_α fixé et pour un ensemble \mathcal{M} de stratégies de recalage les distributions de points $\{(\Delta t(f), \Delta\theta(f))\}_{f \in F}$ mesurés et leur dispersion caractéristique représentée sous forme d'ellipse lors du recalage de différentes séquences². Une

2. Par souci de clarté, nous avons choisi de ne pas inclure dans les graphes de la figure 6.26

bonne stratégie de recalage doit donc amener à observer une petite ellipse proche de l'origine $(0, 0)$ du graphe.

On observe en premier lieu dans la figure 6.26 la forte sensibilité du résultat du recalage séquence sur séquence par rapport au choix du paramètre de tamisage p_α (tests réalisés pour $p_\alpha = 0,7$ et $0,9$), appuyant ainsi l'expérience portant sur l'influence de ce paramètre réalisée en section 6.6.3.3. Si les recalages de la séquence PlantE_P2 sur la séquence PlantE_P1 et de YR_01 sur PlantE_P2 semblent présenter un meilleur comportement pour $p_\alpha = 0,7$, la tendance s'inverse pour les recalages de PlantE_P1 sur PlantE_P2 et surtout de YR_01 sur PlantE_P1. Cela montre de nouveau la nécessité d'appréhender attentivement le problème du paramétrage du recalage spatial robuste.

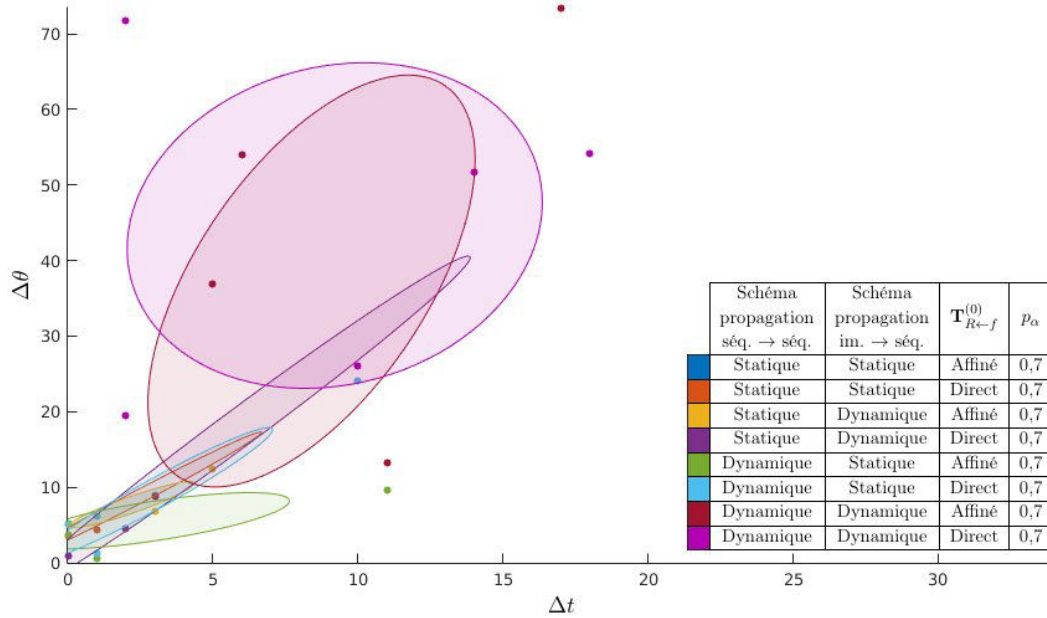
Les mesures réalisées à p_α fixé dévoilent une relative invariance du recalage spatio-temporel des séquences PlantE_P1 sur PlantE_P2 et réciproquement par rapport au choix de la stratégie de propagation $M \in \mathcal{M}$. À l'inverse, le recalage de YR_01 tant sur PlantE_P1 que sur PlantE_P2 montre une importante variabilité du résultat du recalage selon le choix de la stratégie, traduisant une plus forte sensibilité du recalage spatial des images de YR_01 sur ces deux séquences au choix de l'initialisation. Cette sensibilité s'explique en partie par la différence de champ de vue très importante entre les images de YR_01 et les autres séquences de l'étude, occultant ainsi de manière préjudiciable une partie du MF dans les images de YR_01 en vue du recalage de la séquence (se référer à la section 6.6.3.2).

Au regard des résultats présentés en figure 6.26, il est difficile de conclure sur une stratégie « préférentielle » pour la réalisation du recalage spatio-temporel de séquences de méristèmes floraux commune à toutes les séquences. Nous observons toutefois que le paramètre de tamisage p_α approprié varie selon la séquence considérée ($p_\alpha = 0,9$ pour le recalage de PlantE_P1 sur PlantE_P2, $p_\alpha = 0,7$ pour les autres). Considérant p_α fixé à sa valeur appropriée, le choix d'une stratégie basée sur un schéma de propagation de la transformée initiale statique pour le recalage image sur séquence (section 6.5.3), et dynamique pour le recalage séquence sur séquence (section 6.5.4) permet d'obtenir un recalage séquence sur séquence de qualité sinon meilleure, au moins équivalente par rapport aux stratégies basées sur d'autres modes de propagation. Enfin, le choix d'une initialisation de la transformée pour le recalage image sur séquence par la méthode directe permet d'obtenir un recalage de meilleure qualité pour le recalage de YR_01 et de PlantE_P1 sur PlantE_P2, tandis que la méthode affinée apparaît meilleure pour les recalages de YR_01 et PlantE_P2 sur PlantE_P1.

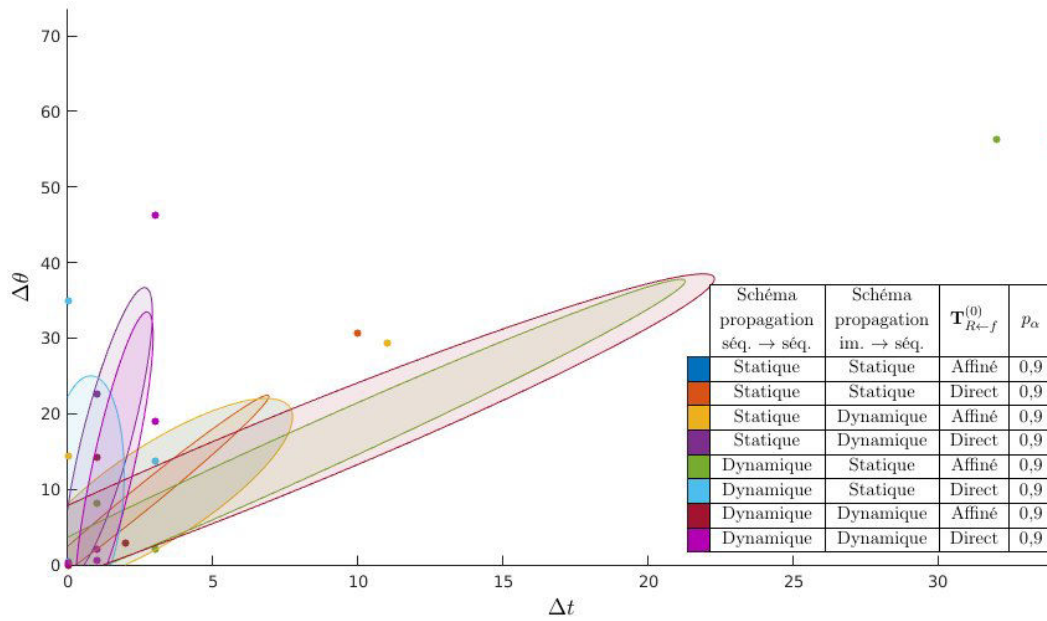
6.7 Conclusions et perspectives

Dans ce chapitre, nous avons proposé une méthode de recalage spatio-temporel de séquences d'images 3D de méristèmes floraux sauvages en développement d'ara-

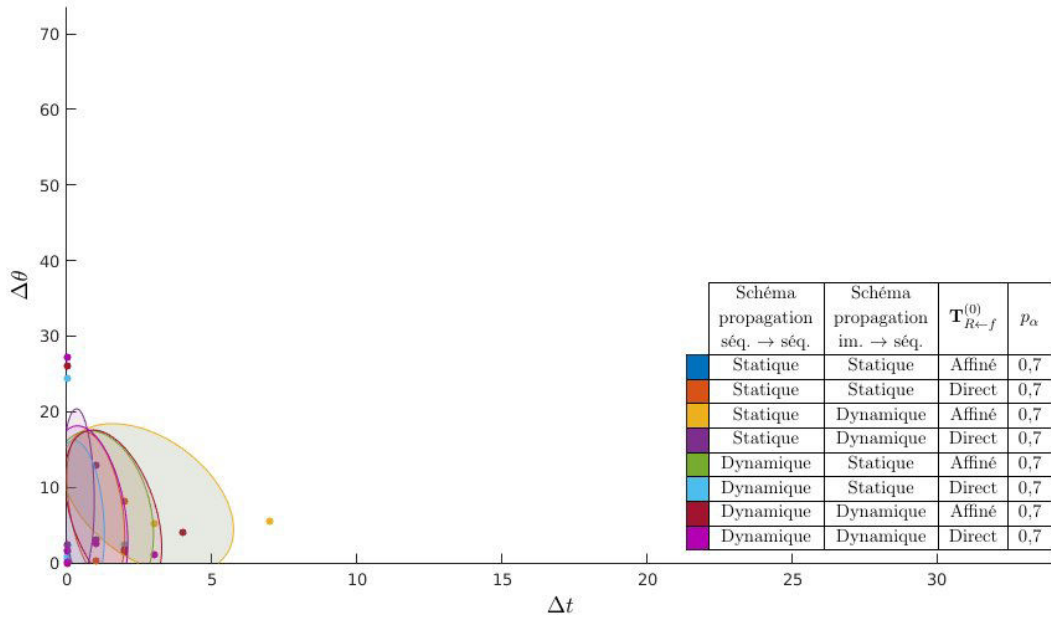
l'information portant sur la translation résiduelle Δ_{trans} mesurée car celle-ci n'enrichissait pas les enseignements pouvant être tirés des mesures réalisées.



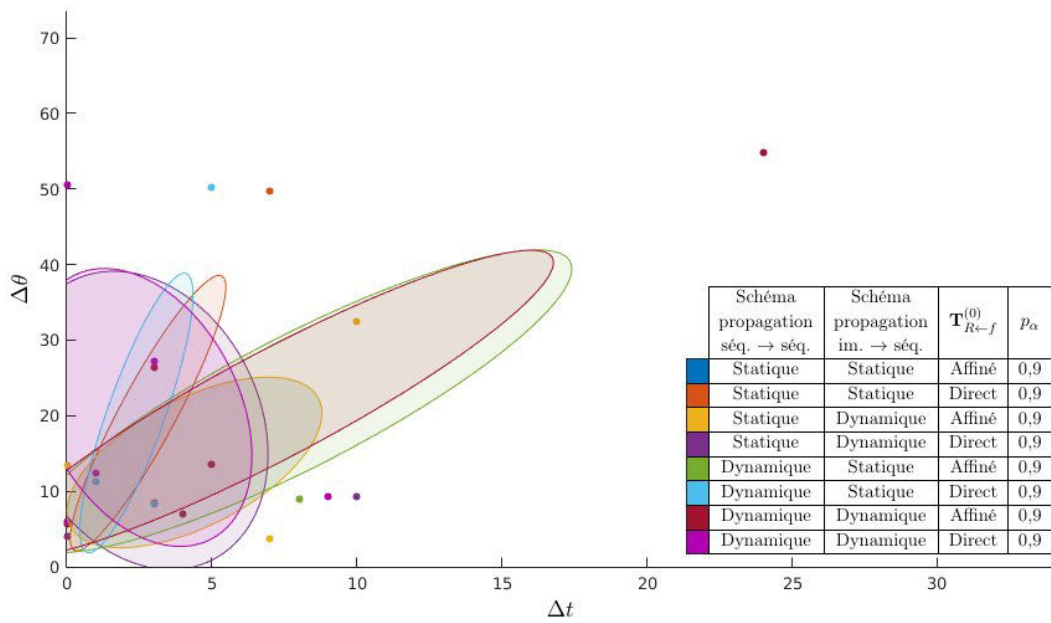
(a) YR_01 recalé sur PlantE_P1 ($p_\alpha = 0,7$)



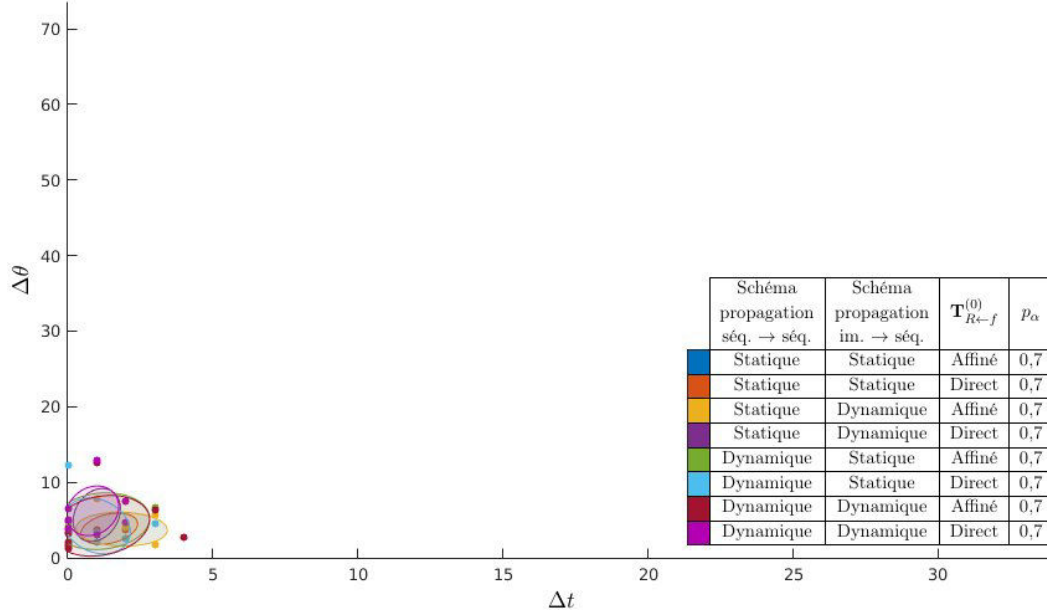
(b) YR_01 recalé sur PlantE_P1 ($p_\alpha = 0,9$)



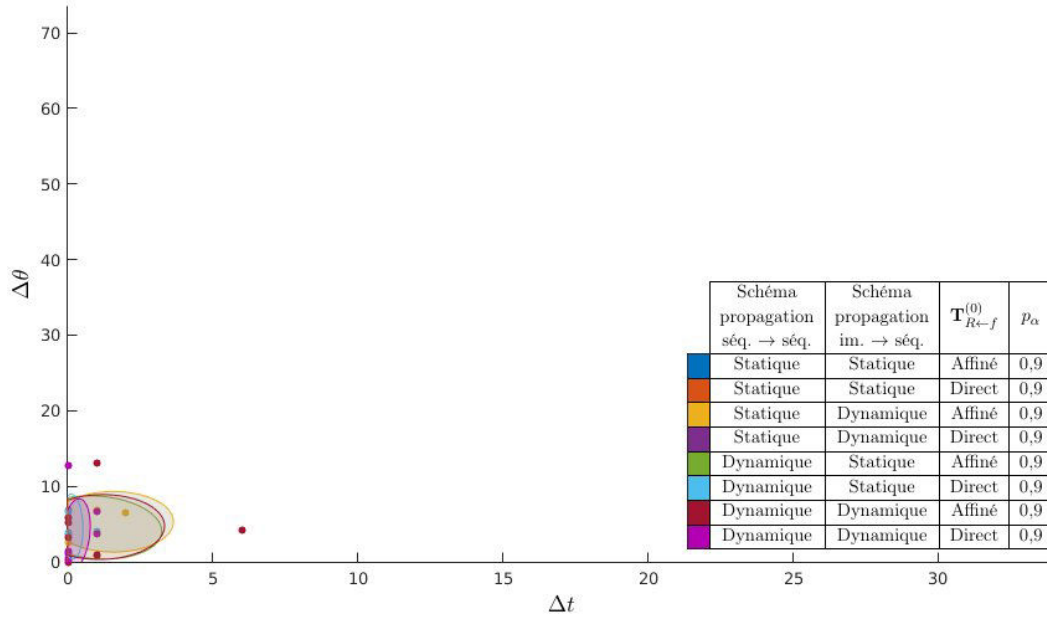
(c) YR_01 recalé sur PlantE_P2 ($p_\alpha = 0,7$)



(d) YR_01 recalé sur PlantE_P2 ($p_\alpha = 0,9$)



(e) PlantE_P1 recalé sur PlantE_P2 ($p_\alpha = 0,7$)



(f) PlantE_P1 recalé sur PlantE_P2 ($p_\alpha = 0,9$)

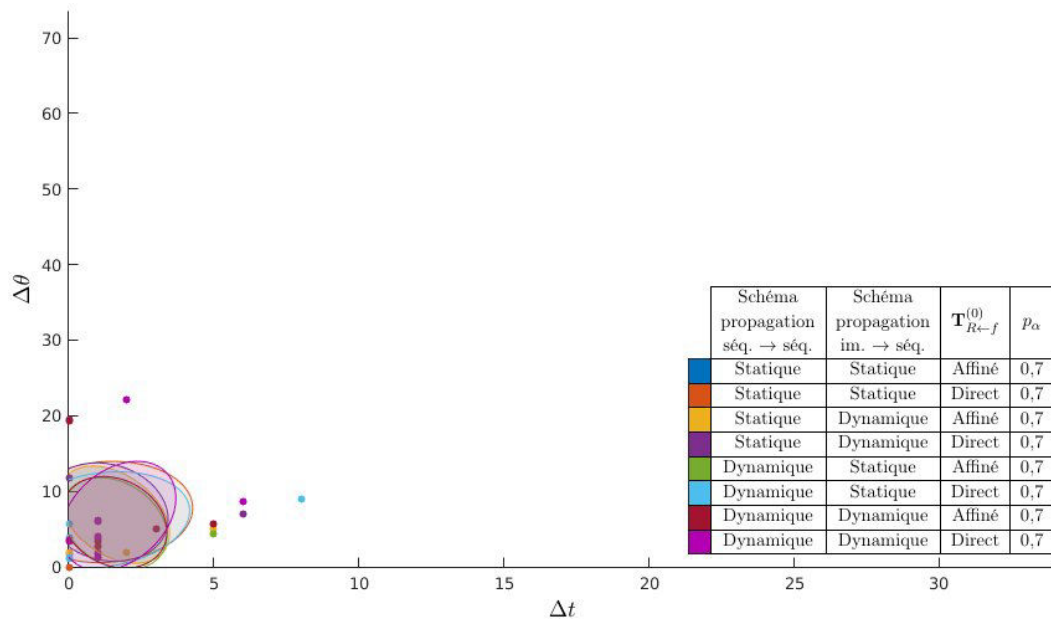
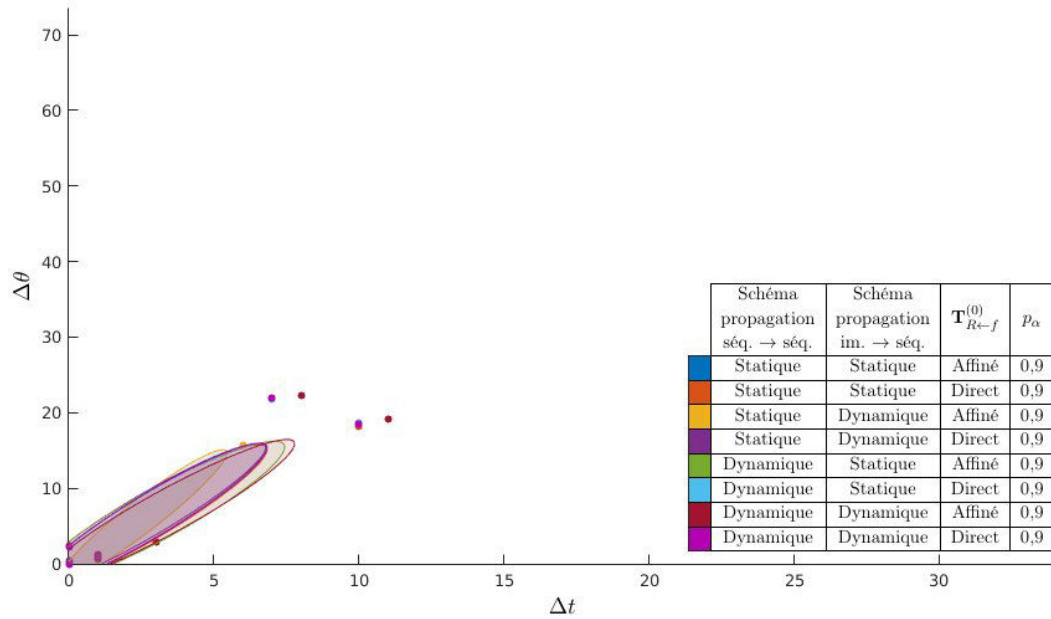
(g) PlantE_P2 recalé sur PlantE_P1 ($p_\alpha = 0,7$)(h) PlantE_P2 recalé sur PlantE_P1 ($p_\alpha = 0,9$)

FIGURE 6.26 – Quantification des différences de temps en heures (en abscisse) et d'angle en degrés (en ordonnée) entre les recalages spatio-temporels de séquences d'images sur des séquences d'images estimés manuellement et ceux obtenus selon plusieurs jeux de paramètres. Chaque point représente les écarts temporel Δt et angulaire $\Delta \theta$ de recalage d'une image de la séquence test sur la séquence cible. La distribution des points obtenus pour un jeu de paramètres donné est représentée par une ellipse caractéristique de la dispersion statistique associée au nuage de points correspondant, qui est mesurée par la matrice de covariance relative à cette distribution. Paramètre d_{\max} fixé à $2 \mu\text{m}$. (a-b) Séquence test YR_01 sur séquence cible PlantE_P1. (c-d) Séquence test YR_01 sur séquence cible PlantE_P2. (e-f) Séquence test PlantE_P1 sur séquence cible PlantE_P2. (g-h) Séquence test PlantE_P2 sur séquence cible PlantE_P1.

bidopsis. Cette méthode exploite les invariances de taille et de forme des MF afin de recalibrer rigide­ment une image de MF « test » sur chacune des images composant la séquence temporelle d’images du MF « de référence ». Un critère de similarité de formes à optimiser permet ensuite d’extraire de la séquence de référence l’instant dont le stade de développement est le plus proche de celui de l’image test. Compte tenu des transformations morphologiques parfois importantes subies par un MF entre deux acquisitions d’images, nous procédons à un sur-échantillonnage de la résolution temporelle de la séquence de référence en calculant des images intermédiaires interpolées à partir de la séquence d’images originale afin d’assurer une détermination précise du recalage temporel. Le recalage d’une séquence test sur une séquence de référence s’obtient par extension du recalage d’une image sur une séquence en considérant une stratégie de propagation de l’initialisation du processus de recalage entre les images successives de la séquence test.

La méthode d’interpolation d’images de MF repose principalement sur la capacité à estimer la transformation non-linéaire (précisément le champ de déformations) subie par un MF entre deux acquisitions d’image. La méthode que nous avons choisie pour estimer cette transformation — le recalage par blocs — est limitée par l’importance des déformations subies par le MF entre les deux acquisitions considérées. Pour lever cette limitation, on peut envisager l’utilisation d’autres outils de recalage (par exemple l’outil MARS-ALT [Fernandez 2010a]). De plus, le choix de la paramétrisation de la méthode de recalage non-linéaire par blocs semble constituer un sujet de réflexion à part entière qui n’a pas été traité dans ce chapitre. Nos expériences ont cependant permis de reconstruire à partir de l’ensemble des séquences dont nous disposons des séquences d’images interpolées convaincantes avec une excellente résolution temporelle.

Nous avons expérimenté la méthode de recalage spatio-temporel décrite dans ce chapitre sur des données réelles provenant de différents MF. Ces expérimentations témoignent de certaines limitations affichées par la méthode actuelle. L’occultation variable des objets à recalibrer, qui résulte d’un déplacement des plantes imagées dans le champ de vue du microscope, constitue une difficulté importante car elle pose la question de la comparabilité des objets à recalibrer. Dans ce contexte, il est indispensable de considérer un recalage spatial robuste. Les expériences portant sur l’influence du coefficient de tamisage de l’algorithme de recalage spatial montrent la sensibilité du recalage au choix du paramétrage de la fonction de coût à optimiser. Il est certainement inadéquat de devoir fixer a priori un paramètre de tamisage pour procéder au recalage spatial de deux MF. Aussi, des stratégies alternatives peuvent être envisagées pour une détermination adaptative plus appropriée du coefficient de tamisage, en prenant par exemple en compte un seuillage sur la distance résiduelle entre les points appariés. Ce seuillage peut par ailleurs être relié au choix du domaine de calcul de la fonction de similarité de formes définie pour l’extraction du recalage temporel.

Des améliorations peuvent également être apportées à la méthode de recalage spatial présentée quant au choix des primitives à recalibrer spatialement. Dans ce chapitre, nous avons fait le choix de considérer des primitives sous la forme de points

et d'apparier les points des MF à recalcr selon un critère de proximité spatiale. Il est possible de considérer d'autres formes de primitives. Par exemple, nous pourrions considérer des primitives de type vecteurs à la place de points, afin d'encoder des informations locales de position, d'épaisseur et de direction normale à la surface. Les appariements de primitives pourraient alors se faire selon un critère plus évolué et la fonction de coût à minimiser pourrait combiner les informations apportées par ce choix de primitives. Il est également envisageable de considérer l'ajout d'un terme de coût supplémentaire basé sur des connaissances a priori sur les images à recalcr, comme le fait que les images de chaque séquence de MF sont deux à deux rigidement recalées. L'extraction de l'axe « vertical » du MF pourrait aussi permettre de contraindre le recalcr spatial de deux méristèmes grâce à l'ajout d'un terme de coût supplémentaire.

La méthode de recalcr temporel présentée dans ce chapitre peut permettre de déterminer la vitesse de développement d'un MF test relativement à une échelle de développement de référence définie par le choix d'un méristème de référence. On peut ainsi comparer les vitesses de développement relatives de différents méristèmes floraux. À partir d'une population de MF temporellement recalés, il est possible d'envisager la construction d'une échelle de développement absolue, c'est-à-dire définie sans qu'il soit nécessaire de choisir une séquence fixant arbitrairement l'échelle de développement de référence.

Bien qu'elle constitue une preuve de concept encourageante, la méthode de recalcr spatio-temporelle que nous avons présentée faillira de manière attendue à recalcr certains spécimens entre eux en raison d'une variabilité morphologique inter-individuelle qui n'est pas toujours négligeable dans les cas extrêmes. La construction d'un multi-atlas agrégeant les données provenant d'une population de MF représentative de la diversité des variétés morphologiques existantes pourrait permettre de lever cette limitation en considérant par la suite un problème de recalcr d'une séquence sur l'ensemble de séquences constituant le multi-atlas. Cette ouverture pose aussi le problème de la quantification de la qualité d'un recalcr spatio-temporel d'une image sur une séquence. Surtout, la construction d'une base de données suffisamment riche pour assurer la représentativité des diverses variétés morphologiques existantes est une opération particulièrement difficile à réaliser d'un point de vue humain et technique (l'acquisition des données et l'obtention de segmentation expertisées pour chaque image sont deux opérations extrêmement chronophages).

Enfin, il serait bien sûr intéressant de voir dans quelle mesure la méthode de recalcr de MF ci-présentée pourrait être étendue au recalcr d'individus mutants, moyennant quelques adaptations. Dans ce contexte, la variabilité de la forme et de la taille des MF est considérablement accrue par rapport au cas de MF sauvages, ce qui invalide les hypothèses d'invariance avec lesquelles nous avons travaillé pour le cas sauvage.

Conclusion générale

Sommaire

7.1 Contributions	139
7.2 Discussions et perspectives	141
7.2.1 Segmentation d'images	141
7.2.2 Recalage d'organismes en développement	141
7.2.3 Mise à disposition des outils développés	144

7.1 Contributions

Le travail réalisé au cours de mes années de thèse a permis de contribuer à apporter des solutions méthodologiques à des problèmes de traitement et d'analyse d'images appliqués à l'étude de la morphogenèse de l'ascidie et de l'arabidopsis.

La numérisation du développement d'organismes en développement passe par la segmentation des cellules qui les composent. Si de nombreux algorithmes de segmentation existent pour réaliser cette tâche [Fernandez 2010a, Mkrtchyan 2011, Mosaliganti 2012, Khan 2014], les erreurs de segmentation sont fréquentes quelle que soit la méthode utilisée. Dans ce contexte, il est important de disposer d'un outil permettant de comparer des segmentations. L'état de l'art fait état de plusieurs outils fournissant des indices de similarité de forme entre des régions correspondantes d'images segmentées. Dans le cas de la segmentation de cellules de tissus épithéliaux, la difficulté réside davantage dans la détection des cellules plutôt que dans l'extraction de leur forme. Comparer deux segmentations d'images de cellules revient dans ce cas à rechercher les différences de détections de cellules. La première contribution apportée par ce manuscrit est la mise en place d'un outil de comparaison de segmentations (chapitre 3) qui vise à identifier les différences de détections de régions ou d'objets segmentés, afin de quantifier les détections « similaires », les sur-détections et les sous-détections d'objets dans une segmentation par rapport à une autre. De plus, des mesures de différences de formes peuvent être extraites entre les régions ainsi mises en correspondance.

Notre seconde contribution (chapitre 4) porte sur la construction d'un outil de segmentation de cellules adapté aux spécificités de l'imagerie par microscopie fluorescente confocale, et en particulier par feuille de lumière. Les tissus que l'on souhaite segmenter sont marqués au niveau des membranes de cellules, formant dans les images des structures brillantes qui entourent le cytoplasme de chaque cellule. Un

algorithme de choix permettant de réaliser la segmentation de ce type d'image est l'algorithme de ligne de partage des eaux [Vincent 1991]. Cependant, les images à traiter sont soumises à des bruits et interruptions de signal qui rendent nécessaire l'application d'un pré-traitement permettant un rehaussement du signal des membranes [Fernandez 2010a, Mosaliganti 2010] que l'on peut conjuguer à une méthode de groupement perceptuel afin de combler les « trous » provoqués par les absences de signal [Mosaliganti 2012]. Notre outil de segmentation se base sur ces principes en réalisant la détection et la reconstruction des membranes cellulaires contenues dans l'image. L'étape de détection est réalisée à l'aide d'un filtre planaire local spécifiquement adapté aux particularités du type d'images traitées. Afin de corriger les inévitables erreurs de sous-détection et de sur-détection de membranes, nous intégrons l'information locale de façon semi-globale à l'aide d'une méthode de groupement perceptuel, le vote de tenseurs. Une attention particulière a été apportée à la complexité algorithmique de l'étape de vote de tenseurs. Nous reconstruisons ainsi une image de membranes adaptée à l'utilisation d'outils morphologiques de segmentation [Vincent 1991] pour l'extraction finale des cellules. Nous utilisons l'outil de comparaison de segmentations développé dans le chapitre 3 pour valider la méthode de segmentation proposée.

Dans les chapitres suivants du manuscrit, notre contribution porte sur la mise en œuvre d'outils de recalage d'individus utiles à des fins de comparaison d'individus [Munro 2006].

Dans le chapitre 5, nous avons exploité le filtre planaire conçu pour l'outil de segmentation de cellules pour orienter les membranes cellulaires chez l'embryon d'ascidie. Nous avons alors montré qu'il est possible d'extraire l'axe de symétrie bilatérale caractéristique de l'embryon d'ascidie à partir de la distribution des orientations spatiales de ses membranes. Nous avons ensuite montré qu'en utilisant cette information de symétrie sous forme de contrainte, il devient possible de recaler spatialement deux embryons d'ascidies à un même stade de développement. Pour cela, nous identifions un ensemble de correspondances cellule à cellule entre les deux embryons. L'algorithme d'optimisation du recalage spatial entre les cellules appariées exploite le caractère fortement stéréotypé du développement des embryons d'ascidies.

Dans le chapitre 6, nous avons exploité des hypothèses d'invariance de taille et de forme des méristèmes floraux de plants d'arabidopsis sauvages conjuguées à la préservation des couches cellulaires épidermique et sub-épidermique des méristèmes apicaux caulinaires pour définir des structures pouvant être recalées rigidement entre les méristèmes floraux. Nous avons montré que sous ces hypothèses, il est également possible de détecter des correspondances temporelles entre des séquences d'images $3D+t$ de méristèmes floraux. En raison des déformations relativement importantes subies par un méristème floral entre deux acquisitions d'images, l'identification de correspondances temporelles entre séquences d'images pourrait être soumise à d'importantes approximations. Afin de répondre à ce problème, nous avons proposé une méthode d'interpolation d'images de méristèmes floraux segmentés. Cette méthode permet d'estimer à n'importe quel instant situé entre deux acquisitions une image segmentée intermédiaire du méristème floral observé. Pour cela, nous nous basons

sur le calcul de la déformation subie par le méristème dans cet intervalle de temps et sur l'identification des lignées cellulaires réalisée à l'aide de la méthode de comparaison de segmentations présentée dans le chapitre 3.

7.2 Discussions et perspectives

7.2.1 Segmentation d'images

La segmentation des cellules composant les organismes en développement est une étape préliminaire indispensable à toute étude quantitative visant à comprendre les phénomènes morphogénétiques. Cependant, il existe un grand nombre de techniques de segmentation d'images et parmi celles spécifiquement conçues pour cette application [Fernandez 2010a, Mosaliganti 2012], aucune ne permet à ce jour de garantir une segmentation « parfaite » des cellules qui composent les tissus imagés. Bien qu'apparaissent de façon régulière de nouvelles techniques performantes, des erreurs de segmentation seront certainement toujours inévitables.

Lors d'analyse de séquences temporelles d'images, il serait certainement plus avisé de considérer le problème de l'extraction de cellules conjointement à celui de l'établissement des lignées cellulaires [Guignard 2015]. Cela revient non-plus à considérer l'objet cellule 3D mais plutôt l'objet lignée cellulaire $3D+t$, mais l'état de l'art montre qu'il demeure difficile d'extraire correctement les lignées cellulaires sur de longues séquences temporelles. Dans le cas de l'ascidie, le caractère stéréotypé de son développement embryonnaire conjugué à la possibilité de recalculer spatialement un embryon modèle supposé parfaitement segmenté — que l'on qualifiera d'*atlas* — sur un individu ouvre de façon intéressante la voie d'une segmentation basée sur le recalage d'*atlas*. Le concept est déjà existant en analyse d'images biomédicales, l'un des standards en la matière étant le référentiel de Talairach [Talairach 1988], système de coordonnées permettant de calquer en tout point du cerveau d'un individu quelconque son point correspondant provenant d'un cerveau-modèle, c'est-à-dire un atlas. La popularité de la segmentation basée sur le recalage d'*atlas* ne cesse d'augmenter (figure 7.1). Nous pensons qu'il serait très avisé d'axer de futures recherches vers cette direction pour la segmentation d'images d'ascidie et l'identification de lignées cellulaires.

7.2.2 Recalage d'organismes en développement

7.2.2.1 Vers la création d'*atlas*

Attardons-nous sur la définition du concept d'*atlas*. Les atlas les plus célèbres, les atlas géographiques, sont des recueils de cartes calquant sur un espace donné représentant une région du monde des informations de natures diverses (géographie, économie, linguistique, etc.). Dans le domaine médical, l'*atlas* de Talairach [Talairach 1988] est un exemple d'*atlas* anatomique représentant un cerveau « modèle » et indiquant la localisation des différentes structures cérébrales qui le composent. Dans le sens large, un atlas se définit ainsi comme étant un espace qui

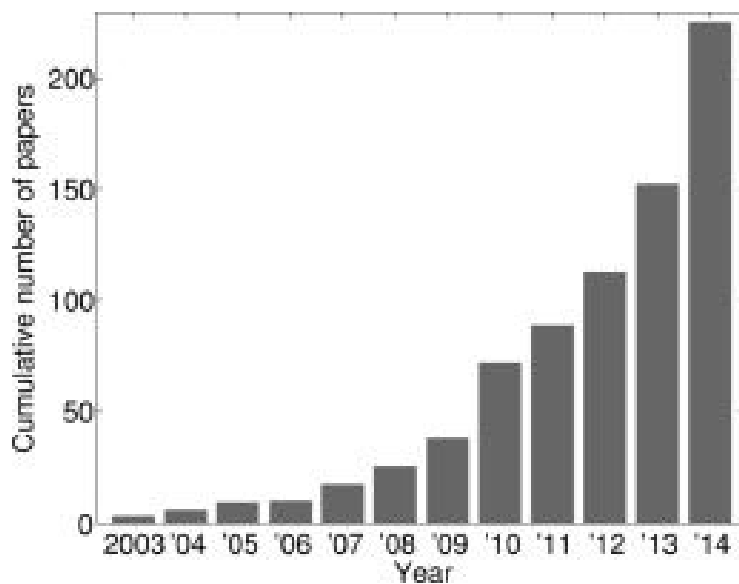


FIGURE 7.1 – Nombre cumulé de papiers introduisant une nouvelle application de segmentation à partir d’atlas multiples référencés dans [Iglesias 2015].

agglomère des connaissances a priori sur une variété d’objets donnée. En biologie développementale, l’atlas peut être défini comme étant un individu représentatif de la variété d’organismes considérée, ou encore une « moyenne » calculée à partir d’un ensemble d’individus de cette variété.

La notion d’atlas de variété suppose que l’on retrouve les mêmes structures d’un individu à l’autre au sein de cette variété et qu’il est possible de projeter ces structures entre elles. Le lien entre la création d’atlas et le problème du recalage d’individus est direct. Il s’agit toujours d’identifier les invariances structurelles, de modéliser le type de transformation permettant de construire des bijections entre ces structures, et de déterminer ces transformations. De façon naturelle, le recalage inter-individus d’organismes en développement amène donc à considérer la question de la création d’atlas, au sens d’images étiquetées, à partir de variétés de ces organismes. Selon l’organisme étudié, les structures invariantes diffèrent.

Dans le cas de l’embryon d’ascidie, il sera ainsi possible d’envisager la construction d’un atlas $3D+t$ représentatif de l’évolution des lignées cellulaires au cours du développement. Dans le cas du méristème floral d’arabidopsis, la conservation des couches cellulaires nous amènerait à considérer la construction d’un atlas $3D+t$ représentant la croissance caractéristique d’un méristème floral à l’échelle des couches cellulaires.

7.2.2.2 Recalage d’embryons

Dans l’optique de la création d’atlas d’embryons, les résultats expérimentaux que nous avons obtenus sur le problème du recalage spatial d’embryons d’ascidies sont encourageants et au demeurant perfectibles. Une première direction de recherche

future sera la considération de l'aspect temporel du recalage d'embryons. Nos expériences montrent que lorsqu'on recalcule une image d'embryon test sur une séquence temporelle de référence provenant d'un autre individu, on observe l'existence d'un intervalle de temps durant lequel le recalage spatial est de bonne qualité. Cet intervalle correspond aux instants de développement de l'embryon de référence où celui-ci possède le même nombre de cellules que l'embryon test.

Si l'on compare le développement de l'embryon d'ascidie avec celui du méristème floral d'arabidopsis, une caractéristique importante de l'embryon d'ascidie est qu'il se développe à volume presque constant. L'identification du stade de développement d'un embryon d'ascidie se fait donc uniquement à partir du dénombrement des cellules qui le composent. Cependant, au cours de sa croissance, un embryon d'ascidie atteint plusieurs paliers de développement au cours desquels, pendant un certain intervalle de temps, le nombre de cellules qui le composent reste constant (stades 32, 44, 64, 76, 112, 128 cellules, etc.). Dès lors, la considération du nombre de cellules en tant que marqueur de temps de développement n'est pas suffisamment spécifique pour permettre une identification précise de ce temps de développement. On peut alors considérer la dynamique temporelle des divisions cellulaires afin de déterminer un temps de développement précis d'un embryon.

Il s'agira donc d'aborder le problème du recalage temporel comme étant un problème de recalage séquence $3D+t$ sur séquence $3D+t$. Parmi les perspectives offertes par la mise au point d'un outil de recalage d'images d'embryons d'ascidies, celle qui a le plus motivé nos travaux est la création d'un atlas $3D+t$, sorte d'individu moyen représentatif de la croissance de l'ascidie et dont le développement est parfaitement connu, c'est-à-dire dont chaque lignée cellulaire est identifiée et nommée. La création d'un tel atlas constituerait une grande avancée dans l'optique phare de conduire des études statistiques portant sur des populations.

La création d'un atlas $3D+t$ décrivant le développement d'un embryon « modèle » numérique à l'échelle de la cellule est soumise à certaines difficultés spécifiques. D'abord, le modèle choisi peut être un individu « moyen » calculé à partir de plusieurs individus segmentés, avec d'éventuelles erreurs de segmentation des cellules. Dans ce cas, l'une des difficultés rencontrées sera de détecter et de corriger ces erreurs en fusionnant des régions dans le cas d'une sur-segmentation ou en les divisant dans le cas d'une sous-segmentation. Une autre difficulté sera la gestion des désynchronisations de divisions cellulaires entre les individus. L'idéal serait de construire un multi-atlas, c'est-à-dire une base d'atlas représentative des variétés développementales d'embryons.

Nous avons enfin ébauché le problème du recalage d'images d'embryons acquises selon différentes modalités, ce qui illustre une autre facette du potentiel de l'outil construit (annexe B).

7.2.2.3 Recalage de méristèmes floraux

La construction d'un outil de recalage spatio-temporel de séquences $3D+t$ d'images de méristèmes floraux constitue notre dernière contribution. Cependant, il serait

présomptueux d'affirmer que l'outil présenté sous sa forme actuelle soit satisfaisant. Les expériences menées montrent en effet l'existence persistante de quelques lacunes dans l'application de notre outil, au premier rang desquelles le choix de la paramétrisation de la méthode constitue un point délicat. Il nous semble que certaines hypothèses de travail peuvent certainement être revues. Notre méthode propose un recalage spatial sous forme de transformation rigide, ce qui est une contrainte probablement trop forte pour permettre de recaler deux à deux les individus de variété morphologique éloignée. Toutefois, les résultats obtenus montrent que dans les cas moins extrêmes, cette hypothèse de travail est raisonnable et conduit à des recalages convaincants de séquences d'images deux à deux. Des difficultés subsistent également dans l'appréhension du recalage des interfaces entre les couches internes de méristèmes floraux.

La suite logique de ce travail serait de considérer désormais un problème de recalage d'une séquence sur une famille de séquences d'images représentative de la variété de morphologies existantes dans la nature. Dans cette logique, nous irions vers la construction d'un multi-atlas $3D+t$ de méristèmes floraux représentatif de la variabilité du développement inter-individuel de ces organismes. Ce multi-atlas pourrait être construit à partir de séquences d'images originales. Il est également envisageable d'enrichir ce multi-atlas en construisant un continuum de séquences $3D+t$ d'images de méristèmes floraux grâce à l'outil d'interpolation d'images décrit dans l'annexe C. Ce faisant, le recalage d'une séquence d'images test sur le multi-atlas pourrait être appréhendé comme un problème d'optimisation alternée de la sélection d'un atlas du multi-atlas morphologiquement voisin de la séquence test et du recalage spatio-temporel de l'atlas sur la séquence.

De même que pour l'embryon d'ascidie, le développement d'un outil de recalage de méristèmes floraux a pour but principal la conduite d'études statistiques inter-individus, et l'un des objectifs futurs sera de construire un outil de comparaison de méristèmes floraux sauvages et mutants. Cela constitue un nouveau défi puisque les hypothèses d'invariance inter-individus formulées dans le cas des plants sauvages ne sont plus vérifiées dans le cas des mutants, avec une plus forte variabilité phénotypique entre individus mutants invalidant la préservation de taille et de forme d'individus à un même stade de développement. Pour s'adapter à ces nouvelles considérations, il sera certainement nécessaire de comprendre la nature de ces variations morphologiques et de considérer un type de transformation qui s'y adapte.

7.2.3 Mise à disposition des outils développés

Pour finir, un important effort reste à faire pour la mise à disposition des différents outils mis en place au cours de cette thèse à la communauté des biologistes. Il s'agirait de greffer ces outils à une plateforme offrant une interface utilisateur familière aux usagers d'outils de traitement d'image. OpenAleaLab est un exemple de plateforme libre multi-paradigme rassemblant des outils variés ayant pour but la modélisation des plantes [Pradal 2008] et sur lequel nous envisageons la création de modules intégrant nos contributions.

Résultats complets de détection des cellules

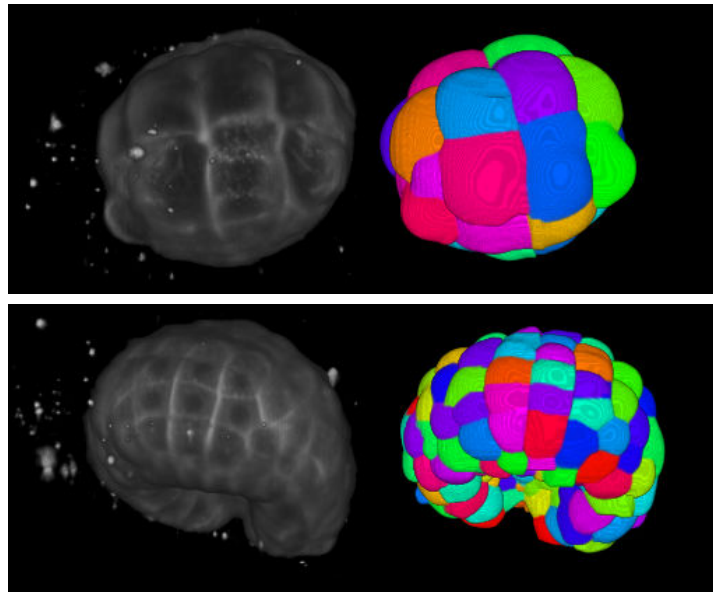


FIGURE A.1 – De gauche à droite : une vue 3D de l'image d'embryon au stade 32 cellules et de sa segmentation (première ligne) ; de même pour l'embryon au stade 162 cellules (deuxième ligne).

Image (#cellules)	Réponse	Vote de tenseurs	σ	h_{\min}	TD	OD	UD	Temps VT (s)		
Im1 (32)	Michelin	BINAIRE	2-6	15	28	4	0	1943		
				30	30	2	0	1943		
		EXTREMA	2-6	10	28	4	0	8253		
				15	28	4	0	8253		
	Mosaliganti	BINAIRE	2	15	26	9	0	2310		
				30	27	9	0			
			3	15	28	6	0	2178		
				30	27	8	0			
			4	15	28	4	0	2220		
				30	29	3	0			
			5	15	28	5	0	2103		
				30	30	3	0			
			6	15	29	9	0	2059		
				30	28	4	0			
			EXTREMA	2	10	17	25	0	22384	
					15	19	20	0		
				3	10	26	8	0	13552	
					15	27	5	0		
				4	10	25	7	0	9268	
					15	24	8	0		
				5	10	24	8	2	6839	
					15	27	5	0		
		6		10	28	4	1	5159		
				15	28	5	0			
		REPONSE		2	10	19	14	0	73722	
					15	26	8	0		
			3	10	28	4	0	65209		
				15	26	7	0			
			4	10	23	9	0	61365		
				15	24	8	0			
			5	10	26	6	1	56516		
				15	26	5	1			
			6	10	25	8	3	42871		
				15	27	4	3			
	MARS [Fernandez 2010a]					31	9	0		

Image (#cellules)	Réponse	Vote de tenseurs	σ	h_{\min}	TD	OD	UD	Temps VT (s)		
Im2 (162)	Michelin	BINAIRE	2-6	15	149	15	3	2422		
				30	145	12	7	2422		
		EXTREMA	2-6	10	144	11	8	7009		
				15	139	15	10	7009		
	Mosaliganti	BINAIRE	2	15	26	9	0	2310		
				30	27	9	0			
			3	15	28	6	0	2178		
				30	27	8	0			
			4	15	28	4	0	2220		
				30	29	3	0			
			5	15	28	5	0	2103		
				30	30	3	0			
			6	15	29	9	0	2059		
				30	28	4	0			
			EXTREMA	2	10	17	25	0	22384	
					15	19	20	0		
		3		10	26	8	0	13552		
				15	27	5	0			
		4		10	25	7	0	9268		
				15	24	8	0			
		5	10	24	8	2	6839			
			15	27	5	0				
		6	10	28	4	1	5159			
			15	28	5	0				
		REPONSE	2	10	19	14	0	73722		
				15	26	8	0			
			3	10	28	4	0	65209		
				15	26	7	0			
	4		10	23	9	0	61365			
			15	24	8	0				
	5		10	26	6	1	56516			
			15	26	5	1				
	6		10	25	8	3	42871			
			15	27	4	3				
	MARS [Fernandez 2010a]					31	9	0		

TABLE A.1 – #cellules : nombre de cellules de la vérité terrain. {Michelin/Mosaliganti} : choix de la fonction de planéité. {REPONSE/EXTREMA/BINAIRE} : procédé utilisé initialiser les primitives du vote de tenseurs (figure 4.16). σ : échelle utilisée pour la fonction de planéité. h_{\min} : paramètre variable de la phase de détection (voir section 4.5). TD/OD/UD : bonnes/sur-/sous-détections. Temps VT : temps mis pour effectuer l'étape de vote de tenseurs (en secondes).

Recalage multimodal d'images d'ascidies

Nous présentons dans cette annexe un premier travail visant à recalibrer des images d'embryons d'ascidies acquises selon des modalités différentes à partir de l'outil de recalage présenté dans le chapitre 5.

B.1 Description des données

Notre expérimentation porte sur le recalage de deux images d'embryons de *Phallusia mammillata* au stade de développement de 64 cellules.

L'image de référence est acquise selon les modalités présentées dans la section 5.5.1. Il s'agit de l'image de l'embryon de référence (Ref. 64) marquée génétiquement au niveau de ses membranes cellulaires et dont la segmentation des cellules a été soumise à expertise. Les 64 barycentres des cellules sont ainsi extraits et constituent un nuage de points $C = \{c_i\}_{i \in [1, 64]}$.

L'image test est marquée au niveau de ses noyaux de cellules, est de taille $512 \times 512 \times 115$ et possède une résolution voxelique de $1 \times 1 \times 2,0234 \mu\text{m}^3$ (figure B.1). L'acquisition de cette image s'effectue *post mortem*, ce qui est à l'origine d'un phénomène observé d'affaissement de l'embryon selon son axe de polarité antéro-postérieur, aligné avec l'axe des z de l'image, ce qui n'affecte pas la symétrie bilatérale de l'embryon. Les barycentres des noyaux sont extraits par un expert, constituant ainsi un nuage de points $D = \{d_i\}_{i \in [1, 64]}$.

B.2 Estimation de l'affaissement de l'embryon test

Le recalage affine du nuage de point D sur C suppose que les embryons dont sont extraits ces nuages de points possèdent une même morphologie, à de faibles variations près. Le phénomène d'affaissement de l'embryon test dû au protocole d'acquisition d'image entrave cette supposition. Avant de procéder au recalage des nuages de points, il est nécessaire d'estimer le coefficient d'affaissement de l'embryon test de manière à estimer les coordonnées qu'auraient les points de D en l'absence d'affaissement.

Nous réalisons une estimation très rudimentaire de ce coefficient en considérant qu'un embryon vivant au stade de 64 cellules possède une forme globale assimilable à une sphère. Nous assimilons l'image de l'embryon test à une boîte englobante de

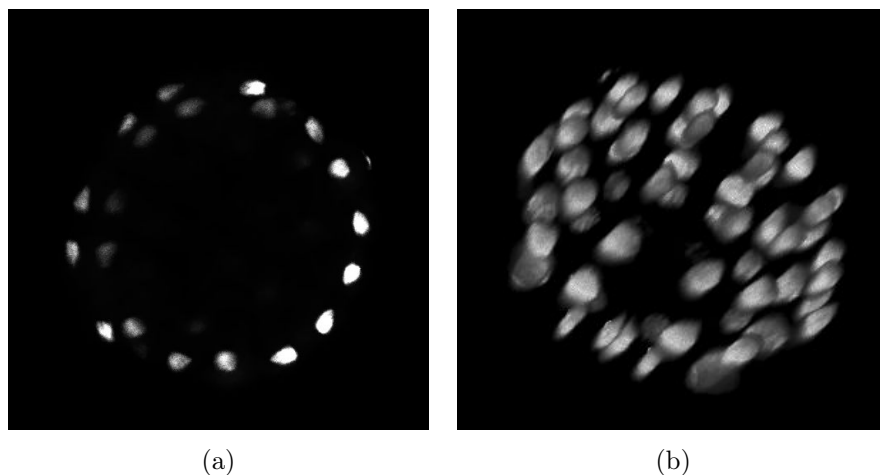


FIGURE B.1 – Image d'embryon d'ascidie obtenue par le marquage des noyaux de cellules. (a) Coupe 2D du volume de données. (b) Vue 3D de l'image.

celui-ci, ce qui constitue une hypothèse réaliste pour cette image. Reconstituer un embryon de forme sphérique revient alors à déterminer le ratio entre les dimensions en $\{x, y\}$ et la dimension en z de l'image test. Le calcul de ce ratio permet de déterminer un coefficient d'affaissement de l'embryon test d'environ 2,2 selon l'axe des z de l'image.

Dans la suite, nous considérons que le nuage de points D est ajusté en tenant compte du coefficient d'affaissement estimé.

B.3 Détection de la symétrie pour l'image test

Dans la section 5.3, nous présentons une méthode automatisée permettant l'extraction de la symétrie bilatérale d'un embryon d'ascidie marqué au niveau de ses membranes cellulaires. Pour extraire la symétrie de l'image test que nous considérons, marquée au niveau des noyaux de cellules, nous avons reconstruit une image de membranes factices à partir du tracé d'un diagramme de Voronoï rogné selon un critère de taille des régions permettant de simuler la forme des cellules de l'embryon à partir des barycentres de cellules (figures B.2a à B.2d). À partir de cette reconstruction, nous appliquons la méthode de la section 5.3 pour l'extraction du plan de symétrie bilatérale de l'embryon test (figures B.2e à B.2h).

Qualitativement, le résultat de l'extraction du plan de symétrie est très satisfaisant pour cette image. Cependant, une étude sur davantage d'individus s'imposerait en vue de valider cette procédure.

B.4 Recalage multimodal

Lors de cette étape, nous appliquons la même méthode de recalage par le calcul d'appariements de barycentres de cellules entre l'embryon de référence et l'embryon

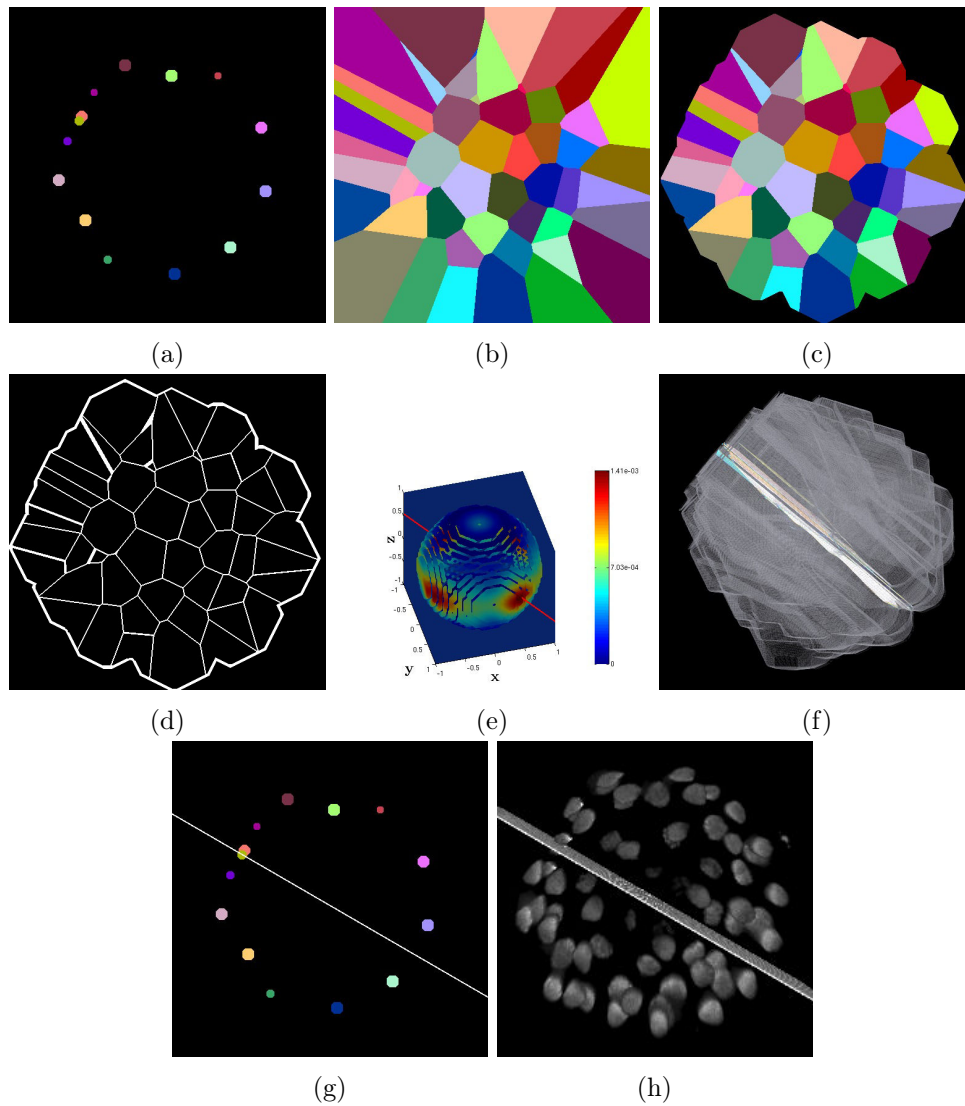


FIGURE B.2 – Étapes de calcul du plan de symétrie bilatérale d'un embryon marqué au niveau des noyaux de cellules. (a) Barycentres dilatés des noyaux de cellules vus en coupe. (b) Pavage de l'image en diagramme de Voronoï construit à partir des barycentres de cellules. (c) Rognage du pavage pour obtenir des reconstructions de cellules de taille réaliste. (d) Extraction des frontières des cellules reconstruites. (e) Estimation par noyau de la densité de probabilité des directions normales associées aux frontières de cellules. (f) Attribution des coefficients de pondération des points de frontières de cellules après sélection du meilleur plan de symétrie candidat. (g) Vue en coupe du plan de symétrie affiné. (h) Vue 3D du plan de symétrie affiné sur l'image d'origine.

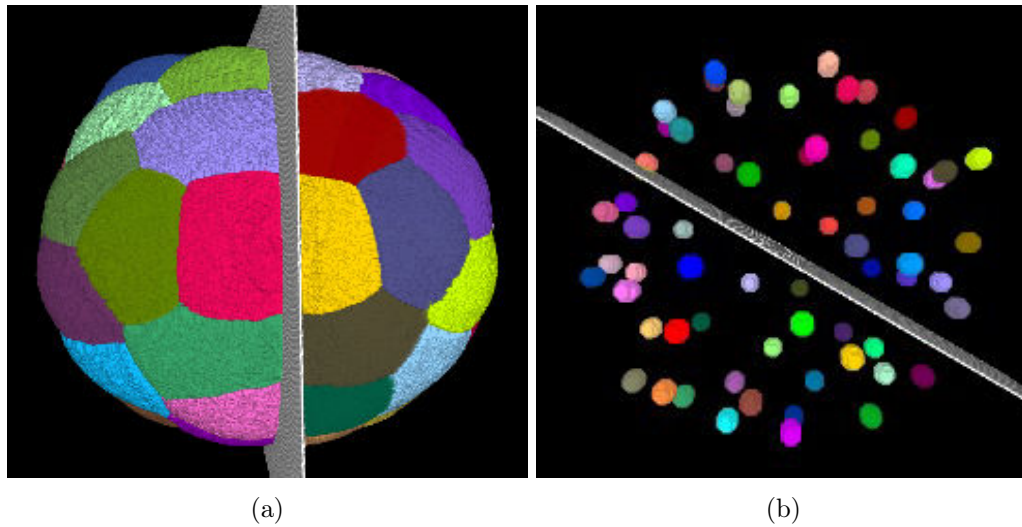


FIGURE B.3 – Vue 3D des données en entrée du processus de recalage multimodal. (a) Embryon de référence : cellules complètes segmentées. (b) Embryon test : barycentres dilatés des noyaux de cellules connus.

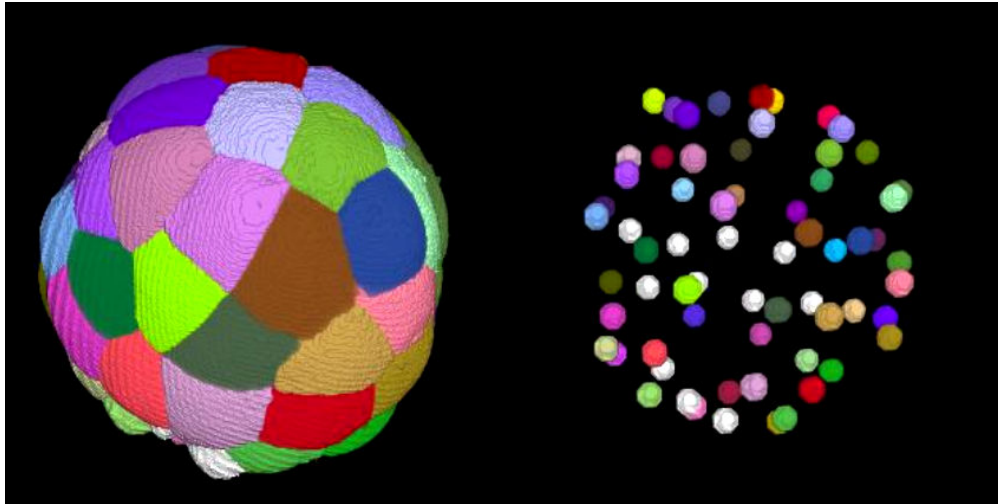
test que celle présentée en section 5.4 et adaptée au recalage d'embryons imagés selon une même méthodologie.

Dans cette annexe, les données que l'on cherche à recaler sont d'une part une image avec des barycentres de cellules complètement segmentées provenant de l'embryon de référence, et d'autre part une image avec des barycentres de noyaux de cellules provenant de l'image test. Le plan de symétrie bilatérale de chacun des embryons est connu (figure B.3).

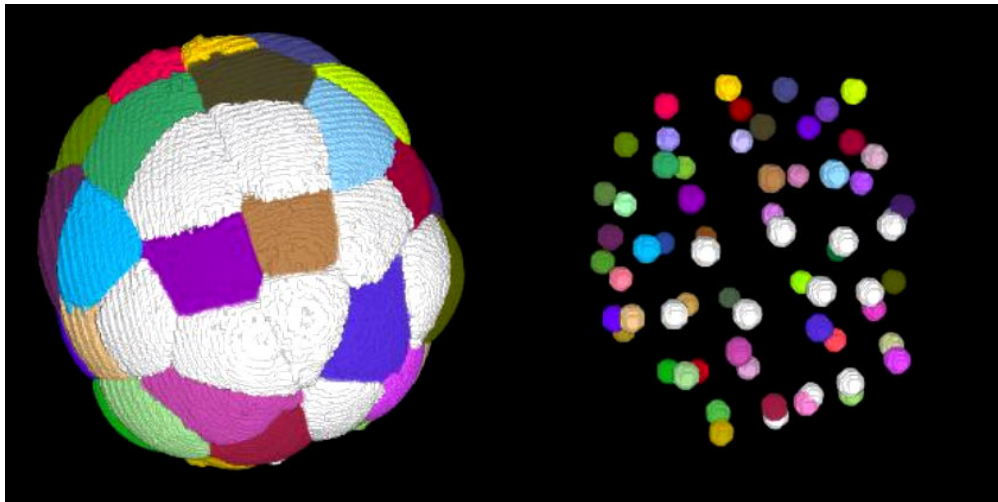
La figure B.4 montre sous deux angles de vue différents les appariements de cellules obtenus par la solution globale du problème de recalage. L'inspection visuelle montre que 78% des cellules des deux embryons sont correctement appariées, alors que ce taux est supérieur à 90% lors du recalage d'embryons imagés selon la même modalité. On dénombre trois erreurs d'appariements, ainsi que 11 cellules non appariées pour chaque embryon sur un total de 64 cellules.

L'investigation de ce résultat montre que beaucoup de cellules de l'embryon de référence mal ou non appariées sont des cellules très allongées (figure B.5). Les appariements sont faits selon le critère du plus proche voisin au sens des distances entre barycentres de cellules et noyaux. Or on observe dans cette figure que les noyaux de l'embryon test peuvent être nettement excentrés par rapport à leur cellule correspondante, et donc distantes de leur barycentre, dans l'embryon de référence.

L'étude de la variation de la distance résiduelle moyenne entre les barycentres des cellules appariées en fonction de la transformation initiale, montrée en figure B.6, soulève un problème de fiabilité du calcul de la solution globale, telle qu'elle est présentée en section 5.4.3, pour le recalage de ces deux images. On observe en effet que si la solution globale du problème de recalage que l'on calcule correspond opportunément à des hypothèses d'appariements de cellules biologiquement sensées,



(a)



(b)

FIGURE B.4 – Appariements de cellules (apparaissant de la même couleur) extraits par la solution globale au problème de recalage, montrés selon deux angles de vue différents. Les cellules non appariées apparaissent en blanc.

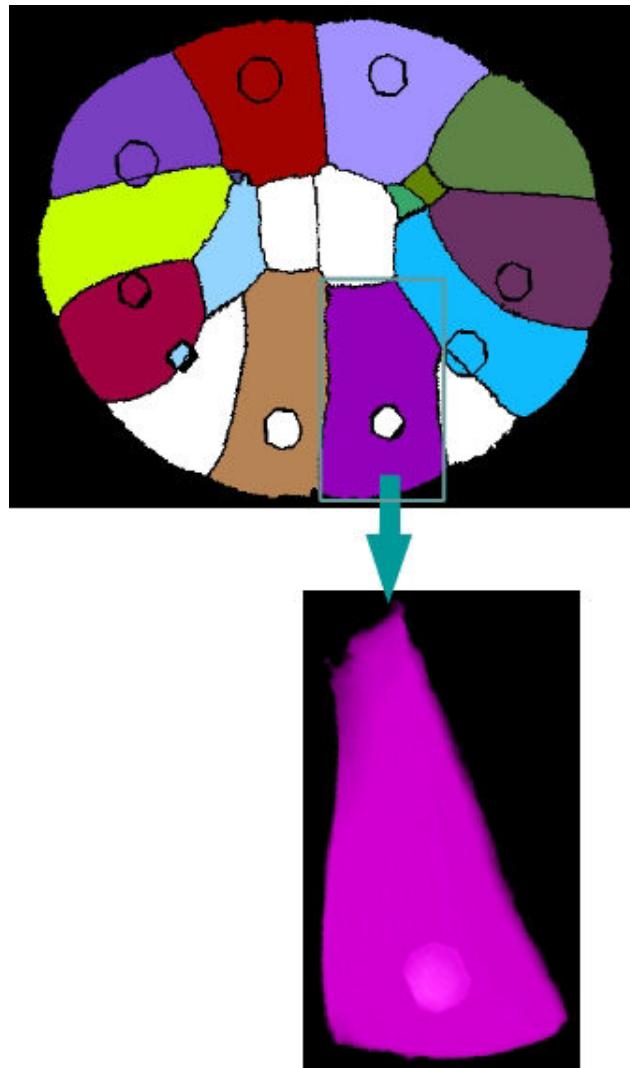


FIGURE B.5 – Vue en coupe de la superposition des deux images recalées. Les régions pleines correspondent aux cellules segmentées de l'embryon de référence. Les régions de forme circulaire correspondent aux barycentres dilatés de l'embryon test. Les cellules de même couleur sont appariées entre elles, tandis que les cellules blanches n'ont pas été mises en correspondance. On représente également une vue 3D de la cellule de référence violette encadrée dans la vue en coupe ainsi que du noyau test blanc qu'elle renferme. L'absence d'appariement entre ces deux cellules constitue une erreur dans la solution globale. Cependant, on peut observer que le noyau blanc est fortement excentré par rapport à la cellule violette, qui possède une forme très allongée. Celle-ci est ainsi par erreur associée au barycentre provenant d'un autre noyau de l'embryon test.

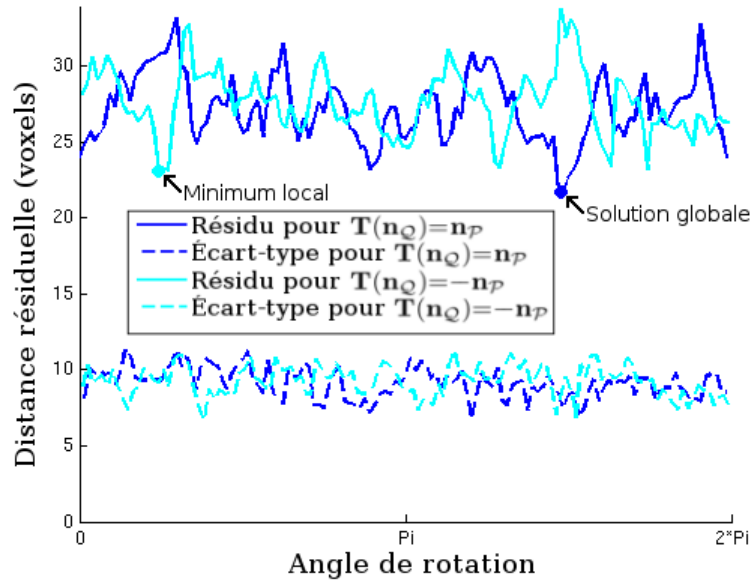


FIGURE B.6 – Variation de la distance résiduelle moyenne entre les barycentres des cellules appariées après optimisation en fonction de l’initialisation du problème de recalage.

un minimum local d’élévation proche de celle de la solution globale apparaît et correspond à des hypothèses d’appariements complètement erronées (figure B.7).

B.5 Conclusion

Nous avons montré dans cette annexe qu’il était possible d’envisager l’identification de correspondances cellule à cellule entre des embryons d’ascidies imagés selon différentes modalités à l’aide de l’outil de recalage développé dans le chapitre 5, moyennant certaines adaptations.

Les résultats que nous présentons montrent cependant quelques lacunes en termes de fiabilité des résultats d’appariements. Ces lacunes sont naturellement expliquées par un certain nombre d’approximations faites au cours de l’expérience de recalage réalisée. Tout d’abord, l’effet d’affaissement de l’embryon imagé selon ses noyaux est pris en compte de façon très rudimentaire. De ce fait, une incertitude importante de nature à faire échouer le recalage s’immisce dans les données de l’embryon test. Enfin, il est assez ambitieux d’envisager que l’hypothèse d’invariance inter-individuelle de la position des barycentres de cellules respectives se vérifie lorsque ceux-ci sont calculés à partir d’une part de cellules entièrement segmentées, et d’autre part de noyaux de cellules seulement.

Bien qu’éminemment perfectible, l’adaptation de la méthode de recalage d’embryons d’ascidies au problème de recalage d’images multimodales que nous avons proposée ici demeure cependant une preuve de concept très encourageante dans la

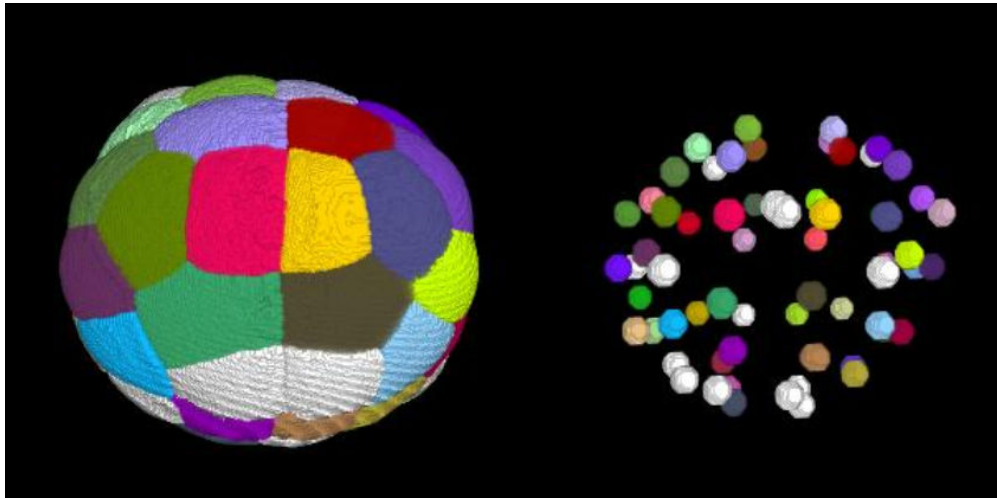


FIGURE B.7 – Appariements de cellules erronés, extraits à partir du minimum local fléché dans la figure B.6.

mesure où ce travail s'appuie sur des hypothèses d'invariance inter-modalités naïves. Nous ne doutons pas qu'en formulant des hypothèses d'invariance plus abouties, l'identification de correspondances cellule à cellule aboutira dans le futur à des résultats convaincants. La capacité de recalibrer à l'échelle cellulaire des embryons imagés selon des modalités diverses constituerait un outil d'intérêt majeur pour l'embryologie.

Interpolation multi-forme

Dans cette annexe, nous considérons le problème de l'interpolation d'images segmentées, c'est-à-dire composées d'une multitude d'objets que l'on qualifera de « formes » dans la suite de cette partie. Ces formes constituent une partition de chacune des images à interpoler et peuvent être sujettes à des déformations.

C.1 Notations

Dans cette annexe, nous considérons deux images I_0 et I_1 et la transformation $\mathbf{T}_{0 \leftarrow 1}$ permettant de ré-échantillonner I_0 sur I_1 , c'est-à-dire de calculer $I_0 \circ \mathbf{T}_{0 \leftarrow 1}$ (éventuellement, $\mathbf{T}_{0 \leftarrow 1} = \text{Id}$ si seule l'interpolation est désirée).

Considérons la partition de I_0 et I_1 respectivement en formes $S_0 = \{s_0^i\}_i$ et $S_1 = \{s_1^j\}_j$. Soit L_0 l'image des labels associée à (I_0, S_0) avec $L_0(M) = \ell_0^i \Leftrightarrow M \in s_0^i$, c'est-à-dire ℓ_0^i est le label associé à la forme s_0^i (respectivement, ℓ_1^j est le label associé à la forme s_1^j).

Un ensemble de correspondances est établi entre des cliques $\{\mathcal{I}_k\}_k$ de labels de L_0 et des cliques $\{\mathcal{J}_k\}_k$ de labels de L_1 . Dire que la clique \mathcal{I}_k correspond à la clique \mathcal{J}_k signifie que la région de I_0 formée par $\cup_{i \in \mathcal{I}_k} s_0^i$ est en correspondance avec la région de I_1 formée par $\cup_{j \in \mathcal{J}_k} s_1^j$. Typiquement, lorsqu'on considère l'étude des lignées cellulaires, \mathcal{I}_k est composé d'une forme (cellule) tandis que \mathcal{J}_k est composé de l'ensemble des formes (cellules) filles de cette dernière. Les correspondances entre cliques peuvent être déterminées en appliquant l'outil de comparaison de segmentations du chapitre 3 entre les partitions $S_0 \circ \mathbf{T}_{0 \leftarrow 1}$ et S_1 .

On dit qu'un couple de labels (ℓ_0^i, ℓ_1^j) est « valide » si ceux-ci appartiennent à des cliques qui se correspondent. Notons \mathcal{V} l'ensemble des couples de labels valides :

$$(\ell_0^i, \ell_1^j) \in \mathcal{V} \Leftrightarrow \exists k | i \in \mathcal{I}_k \text{ et } j \in \mathcal{J}_k. \quad (\text{C.1})$$

Nous considérons que chaque label de chaque image appartient exactement à une clique. Ainsi, pour un label donné $\ell_0^i \in L_0$, nous définissons l'ensemble des labels correspondants $\mathcal{L}(\ell_0^i) \subset L_1$:

$$\mathcal{L}(\ell_0^i) = \{\ell_1^j | (\ell_0^i, \ell_1^j) \in \mathcal{V}\}. \quad (\text{C.2})$$

C.2 Principe

Nous cherchons à estimer des images de labels « compatibles » L_0^t et L_1^t pour $t \in]0, 1[$. Considérons les transformations intermédiaires $\mathbf{T}_{0 \leftarrow t}$ et $\mathbf{T}_{1 \leftarrow t}$ estimées à l'aide

de la méthode décrite dans la section 6.2.3.1. Le point M_t se projette respectivement sur I_0 et I_1 en $M_0 = \mathbf{T}_{0 \leftarrow t}(M_t)$ et $M_1 = \mathbf{T}_{1 \leftarrow t}(M_t)$.

Le principe de la méthode proposée est de trouver le couple valide de labels $(\ell_0^i, \ell_1^j) \in \mathcal{V}$ qui satisfait au mieux une mesure de proximité par rapport aux points M_0 et M_1 (projections de M_t) pondérée en fonction de t . Cette méthode diffère quelque peu de la littérature où l'on trouve des méthodes de calcul de moyennes de formes combinant des segmentations d'un [Herman 1992] ou plusieurs [Rohlfing 2005] objets simultanés afin d'incorporer la connaissance de la transformation $\mathbf{T}_{0 \leftarrow 1}$ ré-échantillonnant I_0 sur I_1 .

Pour tout label ℓ_0^i , nous calculons la distance de M_0 à la forme s_0^i (de même pour les labels ℓ_1^j) :

$$d_0^i = \min_{M | L_0(M) = \ell_0^i} \|MM_0\|. \quad (\text{C.3})$$

Pour chaque couple (ℓ_0^i, ℓ_1^j) , nous calculons un coût $c_{ij} = c(\ell_0^i, \ell_1^j)$ défini par :

$$c_{ij} = (1-t)d_0^i + td_1^j. \quad (\text{C.4})$$

Finalement, L_0^t et L_1^t sont construits par :

$$L_0^t(M_t) = \ell_0^i \text{ et } L_1^t(M_t) = \ell_1^j \text{ avec } (\ell_0^i, \ell_1^j) = \arg \min_{(\ell_0^i, \ell_1^j) \in \mathcal{V}} c_{ij}. \quad (\text{C.5})$$

C.3 Lien avec l'interpolation basée sur la forme (*Shape-based interpolation*)

En interpolation basée sur la forme [Herman 1992], on se situe dans un cadre où une image I est constituée d'une forme S et de son extérieur \bar{S} . L'établissement d'une carte de distances D calculée à partir de l'interface de la forme S , avec une distance (arbitrairement) négative à l'intérieur de la forme S et positive à l'extérieur \bar{S} , permet de redéfinir la frontière de S comme étant l'iso-surface à 0 de la carte de distances (l'objet étant défini par les valeurs de D négatives). À partir des cartes de distances D_0 et D_1 extraites de deux images I_0 et I_1 , on interpole à $t \in]0, 1[$ une image D_t (qui ne constitue pas à proprement parler une image de distances) afin d'interpoler la forme S_t :

$$D_t(M) = (1-t)D_0(M) + tD_1(M). \quad (\text{C.6})$$

La forme S_t est ainsi déduite à partir des valeurs négatives de D_t .

Lorsque $D_0(M)$ et $D_1(M)$ ont un signe opposé, on réalise une interpolation. Considérons arbitrairement le cas où $D_0(M) < 0$ (M est dans S_0) et $D_1(M) > 0$ (M est en dehors de S_1). On a :

$$\begin{cases} M \in S_t & \Leftrightarrow (1-t)D_0(M) + tD_1(M) < 0 & \Leftrightarrow tD_1(M) < -(1-t)D_0(M) \\ M \notin S_t & \Leftrightarrow (1-t)D_0(M) + tD_1(M) > 0 & \Leftrightarrow tD_1(M) > -(1-t)D_0(M) \end{cases} \cdot \quad (\text{C.7})$$

Supposons maintenant que les images considérées sont binaires, avec le label ℓ désignant l'objet et le label $\bar{\ell}$ désignant l'extérieur de l'objet. Si $M \notin S_0$, on a $d_0^{\bar{\ell}} = 0$ (distance nulle séparant le point M de l'extérieur de l'objet) et $d_0^\ell = D_0(M)$ (distance positive à l'extérieur). De même, si $M \in S_1$, on a $d_1^{\bar{\ell}} = 0$ et $d_1^\ell = -D_1(M)$ (distance négative à l'intérieur).

Décider de l'appartenance de M à S_t (label ℓ) ou non (label $\bar{\ell}$) revient à comparer $(1-t)d_0^\ell + td_1^\ell$ et $(1-t)d_0^{\bar{\ell}} + td_1^{\bar{\ell}}$:

$$\begin{aligned} M \in S_t &\Leftrightarrow (1-t)d_0^\ell + td_1^\ell < (1-t)d_0^{\bar{\ell}} + td_1^{\bar{\ell}} \\ &\Leftrightarrow (1-t)d_0^\ell < td_1^{\bar{\ell}} \\ &\Leftrightarrow -(1-t)D_0(M) > tD_1(M). \end{aligned} \tag{C.8}$$

C.4 Détails d'implémentation

Calculer l'ensemble des coûts c_{ij} pour chaque point M_t serait inutile car l'interpolation est uniquement nécessaire au niveau des frontières entre les formes. Le plus souvent en effet, pour des points situés suffisamment à l'intérieur des formes, le couple de labels $(L_0(M_0), L_1(M_1))$ est un couple valide.

Nous proposons une implémentation de la méthode d'interpolation basée sur la recherche de labels à l'intérieur de boules (sphères) $B_{r_0}(M_0)$ et $B_{r_1}(M_1)$ de rayons respectifs r_0 et r_1 centrées en M_0 et M_1 . Nous commençons avec un rayon nul ($r_0 = r_1 = 0$), c'est-à-dire avec le couple $(L_0(M_0), L_1(M_1))$ puis nous augmentons graduellement r_0 et r_1 selon des pas respectifs de δ_0 et δ_1 jusqu'à atteindre un couple valide $(\ell_0^i, \ell_1^j) \in \mathcal{V}$. Afin d'imposer une incrémentation du rayon de recherche de taille au moins égale à un voxel, nous posons :

$$\begin{cases} \text{Si } t < 0,5 & \delta_0 = \frac{t}{1-t} & \delta_1 = 1 \\ \text{Sinon} & \delta_0 = 1 & \delta_1 = \frac{1-t}{t} \end{cases} \cdot \tag{C.9}$$

Supposons que nous avons trouvé $(\ell_0^i, \ell_1^j) \in \mathcal{V}$ avec $\ell_0^i \in L_0 \cap B_{r_0}(M_0)$ et $\ell_1^j \in L_1 \cap B_{r_1}(M_1)$. Le problème consiste à vérifier s'il existe un « meilleur » couple $(\ell_0^{i'}, \ell_1^{j'}) \in \mathcal{V}$ si l'on continue de faire croître r_0 et r_1 , c'est-à-dire vérifiant :

$$(1-t)d_0^{i'} + td_1^{j'} = c_{i'j'} < c_{ij} = (1-t)d_0^i + td_1^j. \tag{C.10}$$

Si un tel couple $(\ell_0^{i'}, \ell_1^{j'})$ existe, il vérifie nécessairement $\ell_0^{i'} \notin L_0 \cap B_{r_0}(M_0)$ ou $\ell_1^{j'} \notin L_1 \cap B_{r_1}(M_1)$. En effet, dans le cas contraire, $(\ell_0^{i'}, \ell_1^{j'})$ aurait déjà été comparé à (ℓ_0^i, ℓ_1^j) .

Sans perte de généralité, supposons que $\ell_0^{i'} \notin L_0 \cap B_{r_0}(M_0)$. On en déduit que $d_0^{i'} > r_0 \geq d_0^i$. L'équation C.10 impose alors que $d_1^{j'} < d_1^j$. On a également $\ell_0^{i'} \in \mathcal{L}(\ell_1^{j'})$.

Une condition nécessaire pour l'existence de $(\ell_0^{i'}, \ell_1^{j'})$ est donc :

$$\begin{aligned} & \exists \ell_0^{i'}, \ell_1^{j'} \text{ tels que} \\ & \text{soit } \left\{ \begin{array}{l} \ell_1^{j'} \in L_1 \cap B_{r_1}(M_1) \mid d_1^{j'} < d_1^j \\ \text{et } \ell_0^{i'} \in \mathcal{L}(\ell_1^{j'}) \text{ avec } \ell_0^{i'} \notin L_0 \cap B_{r_0}(M_0) \end{array} \right. , \\ & \text{soit } \left\{ \begin{array}{l} \ell_0^{i'} \in L_0 \cap B_{r_0}(M_0) \mid d_0^{i'} < d_0^i \\ \text{et } \ell_1^{j'} \in \mathcal{L}(\ell_0^{i'}) \text{ avec } \ell_1^{j'} \notin L_1 \cap B_{r_1}(M_1) \end{array} \right. . \end{aligned} \quad (\text{C.11})$$

Définissons les ensembles I' et J' :

$$\begin{aligned} I' &= \left\{ \ell_0^{i'} \in L_0 \cap B_{r_0}(M_0) \mid d_0^{i'} < d_0^i \text{ et } \exists \ell_1^{j'} \in \mathcal{L}(\ell_0^{i'}) \text{ avec } \ell_1^{j'} \notin L_1 \cap B_{r_1}(M_1) \right\}, \\ J' &= \left\{ \ell_1^{j'} \in L_1 \cap B_{r_1}(M_1) \mid d_1^{j'} < d_1^j \text{ et } \exists \ell_0^{i'} \in \mathcal{L}(\ell_1^{j'}) \text{ avec } \ell_0^{i'} \notin L_0 \cap B_{r_0}(M_0) \right\}. \end{aligned} \quad (\text{C.12})$$

Si les deux ensembles sont vides, alors (ℓ_0^i, ℓ_1^j) est le minimum global. Sinon, plaçons-nous sans perte de généralité dans le cas où $J' \neq \emptyset$. Posons alors $d_1^{j'} = \min_{j' \in J'} d_1^{j'}$. Alors $\ell_0^{i'}$ doit être recherché dans $L_0 \cap B_{R_0}(M_0)$ où $R_0 > r_0$. Une borne pour la valeur de R_0 peut être calculée en considérant l'équation C.10 :

$$\begin{aligned} (1-t)R_0 + td_1^{j'} &\leq (1-t)R_0 + td_1^{j'} < c_{ij} = (1-t)d_0^i + td_1^j \\ \Leftrightarrow R_0 &< \frac{1}{1-t}(c_{ij} - td_1^{j'}) = d_0^i + \frac{t}{1-t}(d_1^j - d_1^{j'}). \end{aligned} \quad (\text{C.13})$$

Épaisseur de la couche L1 du MF d'Arabidopsis

Cette annexe porte sur l'étude de l'épaisseur de la couche de cellules L1 des méristèmes floraux (MF) sauvages. À partir d'un ensemble d'images provenant de cinq individus sauvages (YR_01, PlantE_P1, PlantE_P2, p58 et p194), nous avons extrait les interfaces L0/L1 et L1/L2 de chacune des images segmentées de la base de données (méthode décrite en section 6.3.1).

Considérant une image I aux interfaces extraites et converties en nuages de points $P_{k=1}^I$ pour l'interface L0/L1 et $P_{k=2}^I$ pour l'interface L1/L2, nous avons construit l'histogramme des distances mesurées séparant chaque point de $P_{k=2}^I$ à $P_{k=1}^I$. Pour chaque point de $P_{k=2}^I$, cette mesure a été effectuée en réalisant la recherche du plus proche voisin d'un élément dans un ensemble, à l'aide d'une représentation sous forme d'arbre binaire de recherche k -d de l'ensemble $P_{k=1}^I$ [Bentley 1975].

Les figures D.1 à D.5 présentent les histogrammes de distances ainsi calculés pour les images constituant chacune des séquences temporelles de notre étude. Les histogrammes calculés ont une forme très similaire pour toutes les images, avec un maximum fortement marqué dont l'extraction permet d'estimer une épaisseur « caractéristique » de la L1 sur chaque image.

La figure D.6 permet de visualiser l'évolution au cours du développement de chaque MF de l'épaisseur caractéristique mesurée de la L1. Le premier enseignement de ce graphe est que l'épaisseur caractéristique de la L1 n'est pas constante au cours du développement d'un MF, avec une variation d'épaisseur pouvant atteindre $0,8\mu\text{m}$ (soit environ 20% de l'épaisseur moyenne d'une L1 de MF) pour les séquences PlantE_P1 et PlantE_P2. Cette figure suggère également qu'il existe une importante variation inter-individuelle d'épaisseur de la couche L1 à un même stade de développement. En effet, les instants des différentes séquences ont été manuellement alignés pour rendre compte des correspondances entre les stades de développement de chaque séquence. On observe ainsi une amplitude forte de la variation d'épaisseur de la L1 entre les différents MF pour un même stade de développement.

Cette étude est à replacer dans le contexte du problème de recalage spatial entre MF. La méthode de recalage décrite dans la section 6.3 suppose qu'une transformation rigide permet de superposer les interfaces séparant les couches de cellules L0/L1, L1/L2 et L2/L3 de deux méristèmes floraux à un même stade de développement à de faibles variations de forme près. L'hypothèse semble vérifiée pour l'enveloppe externe du MF, c'est-à-dire pour l'interface L0/L1 (où L0 dénote l'extérieur du méristème).

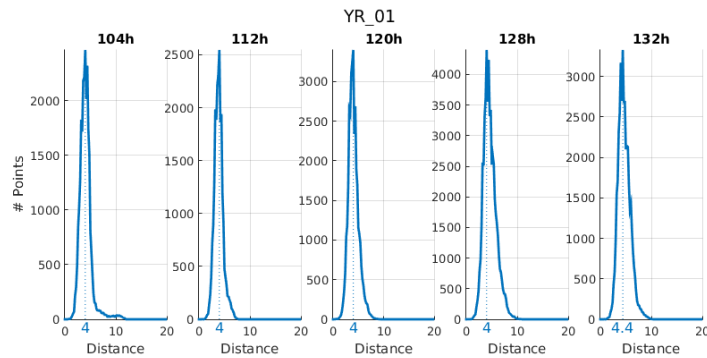


FIGURE D.1 – Histogrammes des distances mesurées séparant les points de l'interface L1/L2 à l'interface L0/L1 sur la séquence d'images YR_01 pour les instants 104h, 112h, 120h, 128h et 132h.

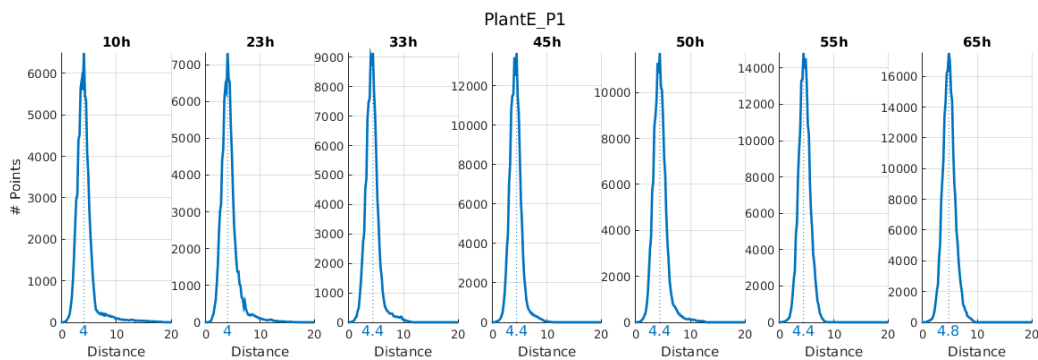


FIGURE D.2 – Histogrammes des distances mesurées séparant les points de l'interface L1/L2 à l'interface L0/L1 sur les images correspondant aux différents instants de la séquence PlantE_P1.

Cependant, l'étude de la présente annexe montre que l'épaisseur de la couche L1 varie selon les individus et au cours du développement de chaque individu. Ainsi, si l'interface L0/L1 peut être considérée comme rigidement invariante à un stade de développement fixe, un facteur de changement d'échelle (c'est-à-dire une similitude) apparaît nécessaire pour recalibrer les interfaces plus profondes. Il n'est donc a priori pas justifié de considérer un recalage rigide entre les interfaces de couches plus profondes (L1/L2 et L2/L3).

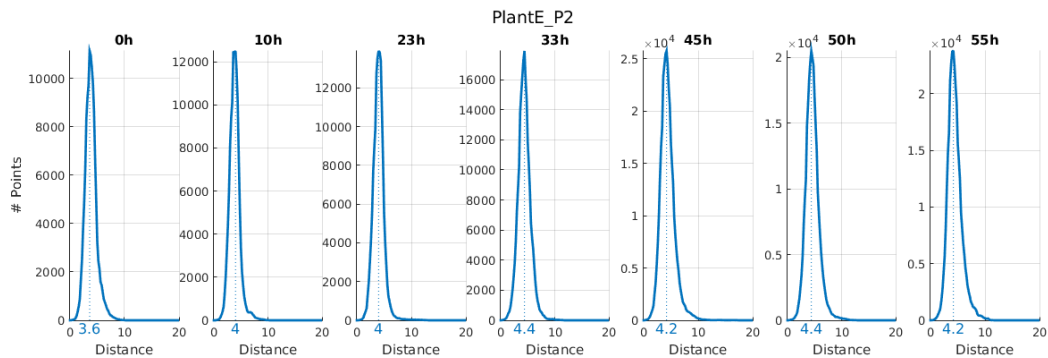


FIGURE D.3 – Histogrammes des distances mesurées séparant les points de l’interface L1/L2 à l’interface L0/L1 sur les images correspondant aux différents instants de la séquence PlantE_P2.

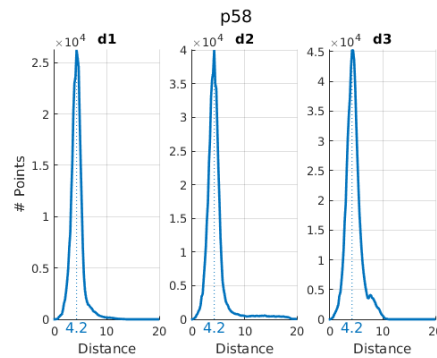


FIGURE D.4 – Histogrammes des distances mesurées séparant les points de l’interface L1/L2 à l’interface L0/L1 sur les images correspondant aux différents instants de la séquence p58.

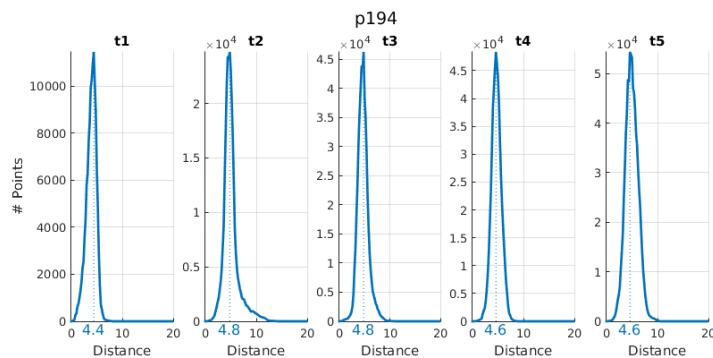


FIGURE D.5 – Histogrammes des distances mesurées séparant les points de l’interface L1/L2 à l’interface L0/L1 sur les images correspondant aux différents instants de la séquence p194.

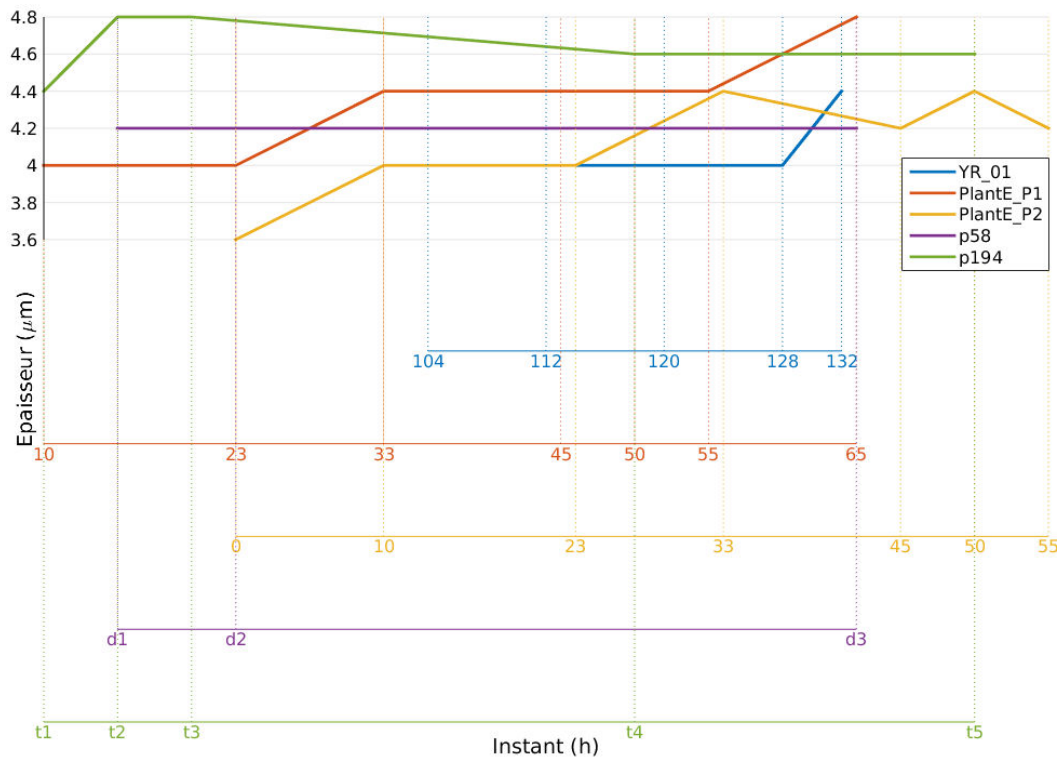


FIGURE D.6 – Visualisation de l'évolution des épaisseurs caractéristiques de couche L1 au cours du développement des différentes séquences d'images de MF. Les échelles de temps des différentes séquences temporelles (i.e. l'instanciation de l'axe des abscisses pour chaque courbe) ont été approximativement alignées afin de rendre compte de la variabilité inter-individu de l'épaisseur de la couche L1 à un même stade de développement.

Bibliographie

- [Adams 1994] R. Adams et L. Bischof. *Seeded region growing*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 6, pages 641–647, Jun 1994. (Cit  en pages 14, 38 et 57.)
- [Al-Kofahi 2006] Omar Al-Kofahi, Richard J Radke, Susan K Goderie, Qin Shen, Sally Temple et Badrinath Roysam. *Automated cell lineage construction : a rapid method to analyze clonal development established with murine neural progenitor cells*. Cell Cycle, vol. 5, no. 3, pages 327–35, February 2006. (Cit  en page 13.)
- [Amat 2014] Fernando Amat, William Lemon, Daniel P Mossing, Katie McDole, Yinan Wan, Kristin Branson, Eugene W Myers et Philipp J Keller. *Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data*. Nat Methods, vol. 11, no. 9, pages 951–8, September 2014. (Cit  en pages 13, 17 et 18.)
- [Ankeny 2011] Rachel A. Ankeny et Sabina Leonelli. *What’s so special about model organisms ?* Studies in History and Philosophy of Science Part A, vol. 42, no. 2, pages 313 – 323, 2011. Model-Based Representation in Scientific Practice. (Cit  en page 4.)
- [Bao 2006] Zhirong Bao, John I Murray, Thomas Boyle, Siew Loon Ooi, Matthew J Sandel et Robert H Waterston. *Automated cell lineage tracing in Caenorhabditis elegans*. Proc Natl Acad Sci U S A, vol. 103, no. 8, pages 2707–12, February 2006. (Cit  en pages 13 et 17.)
- [Barton 1993] MK Barton et RS Poethig. *Formation of the shoot apical meristem in Arabidopsis thaliana : an analysis of development in the wild type and in the shoot meristemless mutant*. Development, vol. 119, no. 3, pages 823–831, 1993. (Cit  en pages 97 et 103.)
- [Bentley 1975] JL Bentley. *Multidimensional binary search trees used for associative searching*. Comm. of the ACM, vol. 18, no. 9, pages 509–517, 1975. (Cit  en pages 107 et 161.)
- [Besl 1992] Paul J. Besl et Neil D. McKay. *A Method for Registration of 3-D Shapes*. IEEE Trans. Pattern Anal. Mach. Intell., vol. 14, no. 2, pages 239–256, F vrier 1992. (Cit  en pages 71, 99 et 106.)
- [Blanchard 2009] Guy B Blanchard, Alexandre J Kabla, Nora L Schultz, Lucy C Butler, Benedicte Sanson, Nicole Gorfinkiel, L Mahadevan et Richard J Adams. *Tissue tectonics : morphogenetic strain rates, cell shape change and intercalation*. Nat Methods, vol. 6, no. 6, pages 458–64, June 2009. (Cit  en page 18.)
- [Bloch 2005] I Bloch, Y Gousseau, H Ma tre, D Matignon, B Pesquet-Popescu, F Schmitt, M Sigelle et F Tupin. *Le traitement des images*. Polycopi  du cours ANIM, version 5.0, vol. 1, 2, 3, 2005. (Cit  en page 12.)

- [Bolker 1995] Jessica A. Bolker. *Model systems in developmental biology*. BioEssays, vol. 17, no. 5, pages 451–455, 1995. (Cité en page 4.)
- [Boulanger 2010] Jérôme Boulanger, Charles Kervrann, Patrick Bouthemy, Peter Elbau, Jean-Baptiste Sibarita et Jean Salamero. *Patch-based nonlocal functional for denoising fluorescence microscopy image sequences*. IEEE transactions on medical imaging, vol. 29, no. 2, pages 442–454, 2010. (Cité en page 12.)
- [Boutet de Monvel 2001] J Boutet de Monvel, S Le Calvez et M Ulfendahl. *Image restoration for confocal microscopy : improving the limits of deconvolution, with application to the visualization of the mammalian hearing organ*. Biophys J, vol. 80, no. 5, pages 2455–70, May 2001. (Cité en pages 12 et 39.)
- [Breiman 2001] Leo Breiman. *Random Forests*. Machine Learning, vol. 45, no. 1, pages 5–32, 2001. (Cité en page 13.)
- [Cachia 2003] A Cachia, J-F Mangin, D Riviere, D Papadopoulos-Orfanos, F Kherif, I Bloch et J Regis. *A generic framework for the parcellation of the cortical surface into gyri using geodesic Voronoi diagrams*. Medical Image Analysis, vol. 7, no. 4, pages 403–16, December 2003. (Cité en pages 57 et 58.)
- [Castro 2009] C Castro, M A Luengo-Oroz, S Desnoullez, L Duloquin, L Fernandez de Manuel, S Montagna, M J Ledesma-Carbayo, P Bourguine, N Peyrieras et A Santos. *An automatic quantification and registration strategy to create a gene expression atlas of zebrafish embryogenesis*. Conf Proc IEEE Eng Med Biol Soc, vol. 2009, pages 1469–72, 2009. (Cité en pages 19 et 70.)
- [Chalfie 1994] M Chalfie, Y Tu, G Euskirchen, W W Ward et D C Prasher. *Green fluorescent protein as a marker for gene expression*. Science, vol. 263, no. 5148, pages 802–5, February 1994. (Cité en page 8.)
- [Cheng 2009] Jierong Cheng et J.C. Rajapakse. *Segmentation of Clustered Nuclei With Shape Markers and Marking Function*. Biomedical Engineering, IEEE Transactions on, vol. 56, no. 3, pages 741–748, March 2009. (Cité en page 38.)
- [Cilla 2015] Rodrigo Cilla, Vinodh Mechery, Beatriz Hernandez de Madrid, Steven Del Signore, Ivan Dotu et Victor Hatini. *Segmentation and tracking of adherens junctions in 3D for the analysis of epithelial tissue morphogenesis*. PLoS Comput Biol, vol. 11, no. 4, April 2015. (Cité en pages 14 et 17.)
- [Darwin 1888] Charles Darwin. *The descent of man, and selection in relation to sex*. J. Murray, 1888. (Cité en page 4.)
- [Day 2009] Richard N Day et Michael W Davidson. *The fluorescent protein palette : tools for cellular imaging*. Chem Soc Rev, vol. 38, no. 10, pages 2887–921, October 2009. Erratum in Chem Soc Rev. 2011 Dec ;40(12) :5923. (Cité en page 11.)
- [Dehal 2002] Paramvir Dehal, Yutaka Satou, Robert K Campbell, Jarrod Chapman, Bernard Degnan, Anthony De Tomaso, Brad Davidson, Anna Di Gregorio, Maarten Gelpke, David M Goodstein, Naoe Harafuji, Kenneth E M Hastings,

- KE Hastings, Isaac Ho, Kohji Hotta, Wayne Huang, Takeshi Kawashima, Patrick Lemaire, Diego Martinez, Ian A Meinertzhagen, Simona Necula, Masaru Nonaka, Nik Putnam, Sam Rash, Hidetoshi Saiga, Masanobu Satake, Astrid Terry, Lixy Yamada, Hong-Gang Wang, Satoko Awazu, Kaoru Azumi, Jeffrey Boore, Margherita Branno, Stephen Chin-Bow, Rosaria DeSantis, Sharon Doyle, Pilar Francino, David N Keys, Shinobu Haga, Hiroko Hayashi, Kyosuke Hino, Kaoru S Imai, Kazuo Inaba, Shungo Kano, Kenji Kobayashi, Mari Kobayashi, Byung-In Lee, Kazuhiro W Makabe, Chitra Manohar et Giorgio Matassi. *The draft genome of *Ciona intestinalis* : insights into chordate and vertebrate origins*. Science, vol. 298, no. 5601, pages 2157–67, December 2002. (Cité en page 5.)
- [Delibaltov 2013] Diana L Delibaltov, Pratim Ghosh, Volkan Rodoplu, Michael Vee-man, William Smith et B S Manjunath. *A linear program formulation for the segmentation of *Ciona* membrane volumes*. In Med Image Comput Comput Assist Interv, volume 8149 of LNCS, pages 444–51. Springer, 2013. (Cité en pages 15 et 38.)
- [Denk 1990] W Denk, J H Strickler et W W Webb. *Two-photon laser scanning fluorescence microscopy*. Science, vol. 248, no. 4951, pages 73–6, April 1990. (Cité en page 10.)
- [Dice 1945] Lee R Dice. *Measures of the amount of ecologic association between species*. Ecology, vol. 26, no. 3, pages 297–302, 1945. (Cité en pages 22 et 78.)
- [Dima 2011] Alden A Dima, John T Elliott, James J Filliben, Michael Halter, Adele Peskin, Javier Bernal, Marcin Kociolek, Mary C Brady, Hai C Tang et Anne L Plant. *Comparison of segmentation algorithms for fluorescence microscopy images of cells*. Cytometry Part A, vol. 79, no. 7, pages 545–559, 2011. (Cité en pages 23, 25 et 27.)
- [Drblíková 2009] Olga Drblíková, Karol Mikula et Nadine Peyriéras. *The Nonlinear Tensor Diffusion in Segmentation of Meaningful Biological Structures from Image Sequences of Zebrafish Embryogenesis*, pages 63–74. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. (Cité en page 41.)
- [Dunham 2006] Margaret H Dunham. *Data mining : Introductory and advanced topics*. Pearson Education India, 2006. (Cité en page 22.)
- [Dzyubachyk 2009] Oleh Dzyubachyk, Rob Jelier, Ben Lehner, Wiros Niessen et Erik Meijering. *Model-based approach for tracking embryogenesis in *Caenorhabditis elegans* fluorescence microscopy data*. Conf Proc IEEE Eng Med Biol Soc, vol. 2009, pages 5356–9, 2009. (Cité en page 13.)
- [Faure 2016] Emmanuel Faure, Thierry Savy, Barbara Rizzi, Camilo Melani, Olga Stasova, Dimitri Fabreges, Robert Spir, Mark Hammons, Robert Cunderlik, Gaelle Recher, Benoit Lombardot, Louise Duloquin, Ingrid Colin, Jozef Kollar, Sophie Desnoulez, Pierre Affaticati, Benoit Maury, Adeline Boyreau, Jean-Yves Nief, Pascal Calvat, Philippe Vernier, Monique Frain, Georges

- Lutfalla, Yannick Kergosien, Pierre Suret, Mariana Remesikova, Rene Dour-sat, Alessandro Sarti, Karol Mikula, Nadine Peyrieras et Paul Bourguine. *A workflow to process 3D+time microscopy images of developing organisms and reconstruct their cell lineage*. Nat Commun, vol. 7, page 8674, 2016. (Cit  en page 13.)
- [Fernandez 2010a] R Fernandez, P Das, V Mirabet, E Moscardi, J Traas, JL Verdeil, G Malandain et C Godin. *Imaging plant growth in 4-D : robust tissue reconstruction and lineaging at cell resolution*. Nat Meth, vol. 7, pages 547–553, 2010. (Cit  en pages 12, 14, 17, 18, 22, 23, 29, 32, 33, 38, 57, 59, 61, 64, 67, 96, 104, 120, 136, 139, 140, 141, 146 et 147.)
- [Fernandez 2010b] Romain Fernandez. *Reconstruction tridimensionnelle et suivi de lign es cellulaires   partir d’images de microscopie laser : application   des tissus v g taux*. Th se de sciences, Universit  Montpellier 2, November 2010. (Cit  en page 17.)
- [Flandin 2002] G Flandin, F Kherif, X Pennec, D Riviere, N Ayache et JB Poline. *Parcellation of brain images with anatomical and functional constraints for fMRI data analysis*. In ISBI, pages 907–910, Washington, United States, 2002. IEEE. (Cit  en page 107.)
- [Fowlkes 2008] Charless C Fowlkes, Cris L Luengo Hendriks, Soile V E Keranen, Gunther H Weber, Oliver Rubel, Min-Yu Huang, Sohail Chatoor, Angela H DePace, Lisa Simirenko, Clara Henriquez, Amy Beaton, Richard Weiszmann, Susan Celniker, Bernd Hamann, David W Knowles, Mark D Biggin, Michael B Eisen et Jitendra Malik. *A quantitative spatiotemporal atlas of gene expression in the Drosophila blastoderm*. Cell, vol. 133, no. 2, pages 364–74, April 2008. (Cit  en pages 19 et 70.)
- [Frangi 1998] AF Frangi, WJ Niessen, KL Vincken et MA Viergever. *Multiscale Vessel Enhancement Filtering*. In MICCAI, volume 1496 of LNCS, pages 130–137. Springer, 1998. (Cit  en page 41.)
- [Frolkovi  2007] Peter Frolkovi , Karol Mikula, Nadine Peyri ras et Alex Sarti. *Counting number of cells and cell segmentation using advection-diffusion equations*. Kybernetika, vol. 43, no. 6, pages 817–829, 2007. (Cit  en page 13.)
- [Gelbart 2012] Michael A Gelbart, Bing He, Adam C Martin, Stephan Y Thiberge, Eric F Wieschaus et Matthias Kaschube. *Volume conservation principle involved in cell lengthening and nucleus movement during tissue morphogenesis*. Proc Natl Acad Sci U S A, vol. 109, no. 47, pages 19298–303, November 2012. (Cit  en page 14.)
- [Guignard 2014] L o Guignard, Christophe Godin, Ulla-Maj Fiuza, Lars Hufnagel, Patrick Lemaire et Gr goire Malandain. *Spatio-temporal registration of embryo images*. In ISBI - International Symposium on Biomedical Imaging, Pekin, Chine, Avril 2014. IEEE. (Cit  en pages 14, 18, 19 et 70.)
- [Guignard 2015] L o Guignard. *Quantitative analysis of animal morphogenesis : from high-throughput laser imaging to 4D virtual embryo in ascidians*. Th se

- de doctorat en science, Université Montpellier, Décembre 2015. (Cité en pages 5, 18, 57, 61, 67 et 141.)
- [Guy 1997] Gideon Guy et Gérard Medioni. *Inference of surfaces, 3D curves, and junctions from sparse, noisy, 3D data*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 19, no. 11, pages 1265–1277, 1997. (Cité en page 51.)
- [Heitger 1994] Friedrich Heitger, Rüdiger von der Heydt et Olaf Kübler. *A computational model of neural contour processing : Figure-ground segregation and illusory contours*. In From Perception to Action Conference, 1994., Proceedings, pages 181–192. IEEE, 1994. (Cité en page 50.)
- [Held 2011] Christian Held, Ralf Palmisano, Lothar Häberle, Michael Hensel et Thomas Wittenberg. *Comparison of parameter-adapted segmentation methods for fluorescence micrographs*. Cytometry Part A, vol. 79, no. 11, pages 933–945, 2011. (Cité en page 23.)
- [Herman 1992] Gabor T Herman, Jingsheng Zheng et C Bucholtz. *Shape-based interpolation*. IEEE Computer Graphics and Applications, vol. 12, no. 3, pages 69–79, 1992. (Cité en pages 103 et 158.)
- [Hotta 2007] Kohji Hotta, Kenta Mitsuhara, Hiroki Takahashi, Kazuo Inaba, Kotaro Oka, Takashi Gojobori et Kazuho Ikeo. *A web-based interactive developmental table for the ascidian *Ciona intestinalis*, including 3D real-image embryo reconstructions : I. From fertilized egg to hatching larva*. Dev Dyn, vol. 236, no. 7, pages 1790–805, July 2007. (Cité en pages 6 et 73.)
- [Huisken 2004] Jan Huisken, Jim Swoger, Filippo Del Bene, Joachim Wittbrodt et Ernst H K Stelzer. *Optical sectioning deep inside live embryos by selective plane illumination microscopy*. Science, vol. 305, no. 5686, pages 1007–9, August 2004. (Cité en page 9.)
- [Iglesias 2015] Juan Eugenio Iglesias et Mert R. Sabuncu. *Multi-atlas segmentation of biomedical images : A survey*. Medical Image Analysis, vol. 24, no. 1, pages 205 – 219, 2015. (Cité en page 142.)
- [Imai 2006] Kaoru S Imai, Michael Levine, Nori Satoh et Yutaka Satou. *Regulatory blueprint for a chordate embryo*. Science, vol. 312, no. 5777, pages 1183–7, May 2006. (Cité en page 5.)
- [Jaccard 1901] Paul Jaccard. Etude comparative de la distribution florale dans une portion des Alpes et du Jura. Impr. Corbaz, 1901. (Cité en page 22.)
- [Johnson 2006] Sam Johnson. *Week-long microscopy course*. http://microscopy.duke.edu/learn/WLC_2.pdf, 2006. (Cité en page 10.)
- [Keller 2008] Philipp J Keller, Annette D Schmidt, Joachim Wittbrodt et Ernst H K Stelzer. *Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy*. Science, vol. 322, no. 5904, pages 1065–9, November 2008. (Cité en pages 12, 13 et 17.)

- [Keller 2010] Philipp J Keller, Annette D Schmidt, Anthony Santella, Khaled Khairy, Zhironng Bao, Joachim Wittbrodt et Ernst H K Stelzer. *Fast, high-contrast imaging of animal development with scanned light sheet-based structured-illumination microscopy*. Nat Methods, vol. 7, no. 8, pages 637–42, August 2010. (Cit  en pages 9 et 13.)
- [Keller 2013] Philipp J. Keller. *Imaging Morphogenesis : Technological Advances and Biological Insights*. Science, vol. 340, no. 6137, pages 1234168+, Juin 2013. (Cit  en pages 8, 12, 38, 69 et 96.)
- [Khan 2014] Zia Khan, Yu-Chiun Wang, Eric F Wieschaus et Matthias Kaschube. *Quantitative 4D analyses of epithelial folding during Drosophila gastrulation*. Development, vol. 141, no. 14, pages 2895–900, July 2014. (Cit  en pages 14, 38 et 139.)
- [Kim 2007] Jong Hyun Kim, Michael S Waterman et Lei M Li. *Diploid genome reconstruction of Ciona intestinalis and comparative analysis with Ciona savignyi*. Genome Res, vol. 17, no. 7, pages 1101–10, July 2007. (Cit  en page 5.)
- [Ko 2009] ByoungChul Ko, MiSuk Seo et Jae-Yeal Nam. *Microscopic cell nuclei segmentation based on adaptive attention window*. Journal of digital imaging, vol. 22, no. 3, pages 259–274, 2009. (Cit  en page 22.)
- [Koffka 1935] Kurt Koffka. Principles of Gestalt psychology. Routledge, 1935. (Cit  en pages 20, 39 et 50.)
- [Korzynska 2007] Anna Korzynska, Wojciech Strojny, Andreas Hoppe, David Wertheim et Pawel Hoser. *Segmentation of microscope images of living cells*. Pattern Analysis and Applications, vol. 10, no. 4, pages 301–319, 2007. (Cit  en page 22.)
- [Kovesi 1997] Peter Kovesi. *Symmetry and Asymmetry from Local Phase*. In Tenth Australian Joint Convergence on Artificial Intelligence, pages 2–4, 1997. (Cit  en page 39.)
- [Krissian 2000] K. Krissian, G. Malandain, N. Ayache, R. Vaillant et Y. Troussel. *Model-Based Detection of Tubular Structures in 3D Images*. Comput Vis Image Underst, vol. 80, no. 2, pages 130–171, 2000. (Cit  en pages 39, 41 et 44.)
- [Kriva 2010] Z Kriva, K Mikula, N Peyrieras, B Rizzi, A Sarti et O Stasova. *3D early embryogenesis image filtering by nonlinear partial differential equations*. Med Image Anal, vol. 14, no. 4, pages 510–26, August 2010. (Cit  en pages 12 et 39.)
- [Krzic 2012] Uros Krzic, Stefan Gunther, Timothy E Saunders, Sebastian J Streichan et Lars Hufnagel. *Multiview light-sheet microscope for rapid in toto imaging*. Nat Methods, vol. 9, no. 7, pages 730–3, July 2012. (Cit  en pages 9, 13, 17, 38, 39, 46, 48, 49, 58 et 83.)
- [Laibach 1943] Frederick Laibach. *Arabidopsis thaliana (L.) Heynh. als Objekt f r genetische und entwicklungsphysiologische Untersuchungen*. Bot Arch, vol. 44, pages 439–455, 1943. (Cit  en page 7.)

- [Landrein 2014] B Landrein, Y Refahi, F Besnard, N Hervieux, V Mirabet, A Boudaoud, T Vernoux et O Hamant. *Meristem size contributes to the robustness of phyllotaxis in Arabidopsis*. Journal of experimental botany, page eru482, 2014. (Cité en pages 97 et 103.)
- [Legrand 2014] J Legrand. *Toward a multi-scale understanding of flower development-from auxin networks to dynamic cellular patterns*. PhD thesis, Lyon, École normale supérieure, 2014. (Cité en pages 97 et 103.)
- [Lemaire 2009] Patrick Lemaire. *Unfolding a chordate developmental program, one cell at a time : invariant cell lineages, short-range inductions and evolutionary plasticity in ascidians*. Dev Biol, vol. 332, no. 1, pages 48–60, August 2009. (Cité en pages 5 et 70.)
- [Li 2009] Fuhai Li, Xiaobo Zhou, Hong Zhao et Stephen TC Wong. *Cell segmentation using front vector flow guided active contours*. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009, pages 609–616. Springer, 2009. (Cité en page 22.)
- [Lindeberg 1994] Tony Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers, Norwell, MA, USA, 1994. (Cité en page 44.)
- [Liu 2014] K. Liu, S. S. Lienkamp, A. Shindo, J. B. Wallingford, G. Walz et O. Ronneberger. *Optical flow guided cell segmentation and tracking in developing tissue*. In IEEE 11th International Symposium on Biomedical Imaging (ISBI), pages 298–301, April 2014. (Cité en page 17.)
- [Liu 2015] Zhe Liu, Luke D Lavis et Eric Betzig. *Imaging live-cell dynamics and structure at the single-molecule level*. Mol Cell, vol. 58, no. 4, pages 644–59, May 2015. (Cité en page 8.)
- [Long 2009] Fuhui Long, Hanchuan Peng, Xiao Liu, Stuart K Kim et Eugene Myers. *A 3D digital atlas of C. elegans and its application to single-cell analyses*. Nat Methods, vol. 6, no. 9, pages 667–72, September 2009. (Cité en pages 12, 13, 19 et 70.)
- [Lorenz 1997] C Lorenz, IC Carlsen, TM Buzug, C Fassnacht et J Weese. *Multi-scale Line Segmentation with Automatic estimation of Width, Contrast and Tangential Direction in 2D and 3D Medical Images*. In CVRMed-MRCAS’97, volume 1205 of LNCS, pages 233–242. Springer, 1997. (Cité en page 41.)
- [Luengo Hendriks 2006] Cris L Luengo Hendriks, Soile V E Keranen, SV Keranen, Charless C Fowlkes, Lisa Simirenko, Gunther H Weber, Angela H DePace, Clara Henriquez, David W Kaszuba, Bernd Hamann, Michael B Eisen, Jitendra Malik, Damir Sudar, Mark D Biggin et David W Knowles. *Three-dimensional morphology and gene expression in the Drosophila blastoderm at cellular resolution I : data acquisition pipeline*. Genome Biol, vol. 7, no. 12, 2006. (Cité en page 19.)
- [Luengo-Oroz 2007] M. A. Luengo-Oroz, E. Faure, B. Lombardot, R. Sance, P. Bourguin, N. Peyrieras et A. Santos. *Twister Segment Morphological Fil-*

- tering. A New Method for Live Zebrafish Embryos Confocal Images Processing.* In IEEE International Conference on Image Processing, volume 5, pages 253–256, Sept 2007. (Cit  en pages 12 et 39.)
- [Luengo-Oroz 2011] M A Luengo-Oroz, M J Ledesma-Carbayo, N Peyrieras et A Santos. *Image analysis for understanding embryo development : a bridge from microscopy to biological insights.* Curr Opin Genet Dev, vol. 21, no. 5, pages 630–7, October 2011. (Cit  en pages 12, 18 et 70.)
- [Maggiori 2014] E. Maggiori, P. Lotito, H. L. Manterola et M. d. Fresno. *Comments on "A Closed-Form Solution to Tensor Voting : Theory and Applications".* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 12, pages 2567–2568, Dec 2014. (Cit  en page 55.)
- [Malandain 2016] Gr goire Malandain. *Blockmatching and associated tools, a very short documentation,* February 2016. (Cit  en page 121.)
- [Malpica 1997] N Malpica, C O de Solorzano, J J Vaquero, A Santos, I Vallcorba, J M Garcia-Sagredo et F del Pozo. *Applying watershed algorithms to the segmentation of clustered nuclei.* Cytometry, vol. 28, no. 4, pages 289–97, August 1997. (Cit  en pages 13 et 38.)
- [Manzanera 2007] A. Manzanera. *Segmentations g om trique et topologique.* http://www.dfr.ensta.fr/Cours/docs/ESE42/Ese4_chap5.pdf, 2007. (Cit  en page 16.)
- [Marcuzzo 2009] Monica Marcuzzo, Tiago Guichard, Pedro Quelhas, Ana Maria Mendon a et Aur lio Campilho. *Cell Division Detection on the Arabidopsis Thaliana Root.* In Helder Araujo, Ana Maria Mendon a, Armando J. Pinho et Mar a In s Torres, editeurs, Pattern Recognition and Image Analysis, pages 168–175, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. (Cit  en pages 17 et 39.)
- [Marr 1980] D Marr et E Hildreth. *Theory of edge detection.* Proc R Soc Lond B Biol Sci, vol. 207, no. 1167, pages 187–217, February 1980. (Cit  en page 13.)
- [Martinez-Sanchez 2013] Antonio Martinez-Sanchez, Inmaculada Garcia et Jose-Jesus Fernandez. *A ridge-based framework for segmentation of 3D electron microscopy datasets.* Journal of Structural Biology, vol. 181, no. 1, pages 61 – 70, 2013. (Cit  en page 47.)
- [Martinez-Sanchez 2014] Antonio Martinez-Sanchez, Inmaculada Garcia, Shoh Asano, Vladan Lucic et Jose-Jesus Fernandez. *Robust membrane detection based on tensor voting for electron tomography.* Journal of Structural Biology, vol. 186, no. 1, pages 49 – 61, 2014. (Cit  en pages 51, 54 et 64.)
- [McMahon 2008] Amy McMahon, Willy Supatto, Scott E Fraser et Angelike Stathopoulos. *Dynamic analyses of Drosophila gastrulation provide insights into collective cell migration.* Science, vol. 322, no. 5907, pages 1546–50, December 2008. (Cit  en page 13.)

- [Medioni 2000] G Medioni, MS Lee et CK Tang. Computational Framework for Segmentation and Grouping. Elsevier Science Inc., New York, NY, USA, 2000. (Cité en pages 39, 51, 52, 53 et 56.)
- [Megason 2007] Sean G. Megason et Scott E. Fraser. *Imaging in Systems Biology*. Cell, vol. 130, no. 5, pages 784–795, 2007. (Cité en pages 3, 19, 69 et 96.)
- [Meyer 1991] Fernand Meyer. *Un algorithme optimal de ligne de partage des eaux*. 8e congrès de reconnaissance des formes et intelligence artificielle, vol. 2, pages 847–859, 1991. (Cité en pages 13 et 14.)
- [Michelin 2014] Gaël Michelin, Léo Guignard, Ulla-Maj Fiuza et Grégoire Maillardain. *Embryo Cell Membranes Reconstruction by Tensor Voting*. In ISBI - International Symposium on Biomedical Imaging, Beijing, China, Avril 2014. IEEE. (Cité en page 47.)
- [Mikula 2011] K Mikula, N Peyrieras, M Remesikova et O Stasova. *Segmentation of 3D cell membrane images by PDE methods and its applications*. Comput Biol Med, vol. 41, no. 6, pages 326–39, June 2011. (Cité en page 15.)
- [Mkrtchyan 2011] K. Mkrtchyan, D. Singh, M. Liu, V. Reddy, A. Roy-Chowdhury et M. Gopi. *Efficient cell segmentation and tracking of developing plant meristem*. In 18th IEEE International Conference on Image Processing, pages 2165–2168, Sept 2011. (Cité en pages 14, 38 et 139.)
- [Moore 2013] Julia L Moore, Zhuo Du et Zhirong Bao. *Systematic quantification of developmental phenotypes at single-cell resolution during embryogenesis*. Development, vol. 140, no. 15, pages 3266–74, August 2013. (Cité en page 18.)
- [Mordohai 2006] Philippos Mordohai et Gérard Medioni. *Tensor voting : a perceptual organization approach to computer vision and machine learning*. Synthesis Lectures on Image, Video, and Multimedia Processing, vol. 2, no. 1, pages 1–136, 2006. (Cité en pages 51 et 56.)
- [Mosaliganti 2010] KR Mosaliganti, F Janoos, A Gelas, RR Noche, N Obholzer, R Machiraju et Sean G Megason. *Anisotropic plate diffusion filtering for detection of cell membranes in 3D microscopy images*. Proc IEEE Int Symp Biomed Imaging, pages 588–591, 2010. (Cité en pages 12, 38, 41 et 140.)
- [Mosaliganti 2012] KR Mosaliganti, RR Noche, F Xiong, IA Swinburne et SG Megason. *ACME : automated cell morphology extractor for comprehensive reconstruction of cell membranes*. PLoS Comput Biol, vol. 8, no. 12, 2012. (Cité en pages 12, 14, 23, 38, 39, 51, 54, 56, 61, 62, 64, 65, 67, 139, 140 et 141.)
- [Mottini 2012] A. Mottini, X. Descombes et F. Besse. *Axon extraction from fluorescent confocal microscopy images*. In 9th IEEE International Symposium on Biomedical Imaging (ISBI), pages 764–767, May 2012. (Cité en page 51.)
- [Munro 2006] Edwin Munro, Francois Robin et Patrick Lemaire. *Cellular morphogenesis in ascidians : how to shape a simple tadpole*. Curr Opin Genet Dev, vol. 16, no. 4, pages 399–405, August 2006. (Cité en pages 37, 69 et 140.)

- [Nishide 2012] Kazuhiko Nishide, Michio Mugitani, Gaku Kumano et Hiroki Nishida. *Neurula rotation determines left-right asymmetry in ascidian tadpole larvae*. *Development*, vol. 139, no. 8, pages 1467–75, April 2012. (Cité en page 72.)
- [Nobelprize.org 2014] Nobelprize.org. *The Nobel Prize in Chemistry 2008*. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2008/, 2014. Nobel Media AB 2014. Web. 29 Apr 2016. (Cité en page 8.)
- [Otsu 1979] N. Otsu. *A threshold selection method from gray-level histograms*. *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pages 62–66, Janvier 1979. (Cité en page 38.)
- [Ourselin 2000] S. Ourselin, A. Roche, S. Prima et N. Ayache. *Block Matching : A General Framework to Improve Robustness of Rigid Registration of Medical Images*. In *Medical Image Computing and Computer-Assisted Intervention*, volume 1935 of *LNCIS*, pages 557–566, Berlin, Heidelberg, 2000. Springer. (Cité en pages 17, 99, 100 et 121.)
- [Parent 1989] Pierre Parent et Steven W Zucker. *Trace inference, curvature consistency, and curve detection*. *Pattern Analysis and Machine Intelligence*, *IEEE Transactions on*, vol. 11, no. 8, pages 823–839, 1989. (Cité en pages 50 et 51.)
- [Parzen 1962] Emanuel Parzen. *On estimation of a probability density function and mode*. *The annals of mathematical statistics*, vol. 33, no. 3, pages 1065–1076, 1962. (Cité en page 76.)
- [Polak 2009] Mark Polak, Hong Zhang et Minghong Pi. *An evaluation metric for image segmentation of multiple objects*. *Image and Vision Computing*, vol. 27, no. 8, pages 1223–1227, 2009. (Cité en page 22.)
- [Pons 2005] Thomas Pons et Jérôme Mertz. *Imager dans la profondeur des tissus : les microscopies biphotoniques*. *Images de la physique*, pages 159–165, 2005. (Cité en pages 10 et 11.)
- [Pop 2011] S Pop, A Dufour et JC Olivo-Marin. *Image filtering using anisotropic structure tensor for cell membrane enhancement in 3D microscopy*. In *Image Processing (ICIP)*, 18th IEEE International Conference on, pages 2041–2044, Sept 2011. (Cité en page 41.)
- [Pradal 2008] Christophe Pradal, Samuel Dufour-Kowalski, Frédéric Boudon, Christian Fournier et Christophe Godin. *OpenAlea : a visual programming and component-based software platform for plant modelling*. *Functional plant biology*, vol. 35, no. 10, pages 751–760, 2008. (Cité en page 144.)
- [Provart 2016] Nicholas J Provart, Jose Alonso, Sarah M Assmann, Dominique Bergmann, Siobhan M Brady, Jelena Brkljacic, John Browse, Clint Chapple, Vincent Colot, Sean Cutler, Jeff Dangl, David Ehrhardt, Joanna D Friesner, Wolf B Frommer, Erich Grotewold, Elliot Meyerowitz, Jennifer Nemhauser, Magnus Nordborg, Craig Pikaard, John Shanklin, Chris Somerville, Mark

- Stitt, Keiko U Torii, Jamie Waese, Doris Wagner et Peter McCourt. *50 years of Arabidopsis research : highlights and future directions*. New Phytol, vol. 209, no. 3, pages 921–44, February 2016. (Cit  en page 7.)
- [Qi 2014] J Qi. *Dense nuclei segmentation based on graph cut and convexity-concavity analysis*. J Microsc, vol. 253, no. 1, pages 42–53, January 2014. (Cit  en page 38.)
- [Rey 2002] David Rey. *Automatic detection and quantification of evolving processes in 3D medical images : application to multiple sclerosis*. Theses, Universit  Nice Sophia Antipolis, Octobre 2002. (Cit  en page 102.)
- [Risser 2008] L. Risser, F. Plouraboue et X. Descombes. *Gap Filling of 3-D Microvascular Networks by Tensor Voting*. IEEE Transactions on Medical Imaging, vol. 27, no. 5, pages 674–687, May 2008. (Cit  en page 51.)
- [Rohlfing 2005] Torsten Rohlfing et Calvin R Maurer Jr. *Shape-based averaging for combination of multiple segmentations*. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 838–845. Springer, 2005. (Cit  en page 158.)
- [Rousseeuw 1987] PJ Rousseeuw et AM Leroy. *Robust Regression and Outlier Detection*. John Wiley & Sons, New York, NY, USA, 1987. (Cit  en pages 82 et 106.)
- [Russell 2009] Richard A Russell, Niall M Adams, David A Stephens, Elizabeth Batty, Kirsten Jensen et Paul S Freemont. *Segmentation of fluorescence microscopy images for quantitative analysis of cell nuclear architecture*. Biophysical journal, vol. 96, no. 8, pages 3379–3389, 2009. (Cit  en page 22.)
- [Sander 1986] Peter T Sander et Steven W Zucker. *Stable surface estimation*. In Proceedings of the Eighth International Conference on Pattern Recognition, Paris, France, pages 1165–1167, 1986. (Cit  en page 51.)
- [Santella 2010] Anthony Santella, Zhuo Du, Sonja Nowotschin, Anna-Katerina Hadjantonakis et Zhirong Bao. *A hybrid blob-slice model for accurate and efficient detection of fluorescence labeled nuclei in 3D*. BMC Bioinformatics, vol. 11, page 580, 2010. (Cit  en page 18.)
- [Sarder 2006] Pinaki Sarder et Arye Nehorai. *Deconvolution methods for 3-D fluorescence microscopy images*. Signal Processing Magazine, IEEE, vol. 23, no. 3, pages 32–45, 2006. (Cit  en page 13.)
- [Sarkar 1994] Sudeep Sarkar et Kim L Boyer. *A computational structure for preattentive perceptual organization : Graphical enumeration and voting methods*. Systems, Man and Cybernetics, IEEE Transactions on, vol. 24, no. 2, pages 246–267, 1994. (Cit  en page 51.)
- [Sarti 2002] A. Sarti, R. Malladi et J.A. Sethian. *Subjective Surfaces : A Geometric Model for Boundary Completion*. International Journal of Computer Vision, vol. 46, no. 3, pages 201–221, 2002. (Cit  en pages 15 et 38.)

- [Schindelin 2012] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid *et al.* *Fiji : an open-source platform for biological-image analysis*. *Nature methods*, vol. 9, no. 7, pages 676–682, 2012. (Cité en pages 106 et 115.)
- [Schultz 1991] Elizabeth A Schultz et George W Haughn. *LEAFY, a homeotic gene that regulates inflorescence development in Arabidopsis*. *The Plant Cell*, vol. 3, no. 8, pages 771–781, 1991. (Cité en page 96.)
- [Shapiro 2015] Bruce E Shapiro, Cory Tobin, Eric Mjolsness et Elliot M Meyerowitz. *Analysis of cell division patterns in the Arabidopsis shoot apical meristem*. *Proceedings of the National Academy of Sciences*, vol. 112, no. 15, pages 4815–4820, 2015. (Cité en page 102.)
- [Soille 1999] P Soille. *Morphological image analysis : principles and applications*. Springer, 1999. (Cité en page 58.)
- [Sommer 2011] C. Sommer, C. Straehle, U. Köthe et F. A. Hamprecht. *Ilastik : Interactive learning and segmentation toolkit*. In *IEEE International Symposium on Biomedical Imaging : From Nano to Macro*, pages 230–233, March 2011. (Cité en page 13.)
- [Sotiras 2013] Aristeidis Sotiras, Christos Davatzikos et Nikos Paragios. *Deformable medical image registration : a survey*. *IEEE Trans Med Imaging*, vol. 32, no. 7, pages 1153–90, July 2013. (Cité en pages 17, 19 et 71.)
- [Srinivasa 2009] Gowri Srinivasa, Matthew C Fickus, Yusong Guo, Adam D Linstead et Jelena Kovačević. *Active mask segmentation of fluorescence microscope images*. *Image Processing, IEEE Transactions on*, vol. 18, no. 8, pages 1817–1829, 2009. (Cité en page 22.)
- [Stegmaier 2016] Johannes Stegmaier, Fernando Amat, William C Lemon, Katie McDole, Yinan Wan, George Teodoro, Ralf Mikut et Philipp J Keller. *Real-Time Three-Dimensional Cell Segmentation in Large-Scale Microscopy Data of Developing Embryos*. *Dev Cell*, vol. 36, no. 2, pages 225–40, January 2016. (Cité en pages 15 et 38.)
- [Stolfi 2014] Alberto Stolfi, Elijah K Lowe, Claudia Racioppi, Filomena Ristoratore, C Titus Brown, Billie J Swalla et Lionel Christiaen. *Divergent mechanisms regulate conserved cardiopharyngeal development and gene expression in distantly related ascidians*. *Elife*, vol. 3, 2014. (Cité en page 5.)
- [Supatto 2009] Willy Supatto, Amy McMahon, Scott E Fraser et Angelike Stathopoulos. *Quantitative imaging of collective cell migration during Drosophila gastrulation : multiphoton microscopy and computational analysis*. *Nat Protoc*, vol. 4, no. 10, pages 1397–412, 2009. (Cité en page 13.)
- [Surkova 2013] Svetlana Surkova, Ekaterina Myasnikova, Konstantin N Kozlov, Andrei Pisarev, John Reinitz et Maria Samsonova. *Quantitative imaging of gene expression in Drosophila embryos*. *Cold Spring Harb Protoc*, vol. 2013, no. 6, pages 488–97, June 2013. (Cité en page 19.)

- [Talairach 1988] J. Talairach et P. Tournoux. Co-planar Stereotaxic Atlas of the Human Brain : 3-dimensional Proportional System : an Approach to Cerebral Imaging. Thieme classics. G. Thieme, 1988. (Cit  en page 141.)
- [Tassy 2006] Olivier Tassy, Fabrice Daian, Clare Hudson, Vincent Bertrand et Patrick Lemaire. *A quantitative approach to the study of cell shapes and interactions during early chordate embryogenesis*. Curr Biol, vol. 16, no. 4, pages 345–58, February 2006. (Cit  en page 18.)
- [Tong 2004] Wai-Shun Tong, Chi-Keung Tang, P. Mordohai et G. Medioni. *First order augmentation to tensor voting for boundary inference and multiscale analysis in 3D*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 5, pages 594–611, May 2004. (Cit  en page 51.)
- [Traas 2009] Jan Traas et Olivier Hamant. *From genes to shape : understanding the control of morphogenesis at the shoot meristem in higher plants using systems biology*. C R Biol, vol. 332, no. 11, pages 974–85, November 2009. (Cit  en page 37.)
- [Ullman 1988] Shimon Ullman et Amnon Sha’ashua. *Structural saliency : The detection of globally salient structures using a locally connected network*. Proceedings of International Conference on Computer Vision, 1988. (Cit  en page 50.)
- [Vincent 1991] L Vincent et P Soille. *Watersheds in Digital Spaces : An Efficient Algorithm Based on Immersion Simulations*. IEEE Trans Pattern Anal Mach Intell, vol. 13, no. 6, pages 583–598, 1991. (Cit  en pages 58 et 140.)
- [Waterwereld 2002] Waterwereld. *Vase tunicate or Ciona intestinalis*. <http://www.waterwereld.nu/vase.php>, 2002. (Cit  en page 5.)
- [Weber 2014] Michael Weber, Michaela Mickoleit et Jan Huisken. *Light sheet microscopy*. Methods Cell Biol, vol. 123, pages 193–215, 2014. (Cit  en page 10.)
- [Wikipedia 2015] Wikipedia. *Ligne de partage des eaux (segmentation)*. [https://fr.wikipedia.org/wiki/Ligne_de_partage_des_eaux_\(segmentation\)](https://fr.wikipedia.org/wiki/Ligne_de_partage_des_eaux_(segmentation)), 2015. (Cit  en page 16.)
- [Wu 2012] T. P. Wu, S. K. Yeung, J. Jia, C. K. Tang et G. Medioni. *A Closed-Form Solution to Tensor Voting : Theory and Applications*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 8, pages 1482–1495, Aug 2012. (Cit  en page 55.)
- [Zanella 2010] Cecilia Zanella, Matteo Campana, Barbara Rizzi, Camilo Melani, Gonzalo Sanguinetti, Paul Bourguine, Karol Mikula, Nadine Peyrieras et Alessandro Sarti. *Cells segmentation from 3-D confocal images of early zebrafish embryogenesis*. IEEE Trans Image Process, vol. 19, no. 3, pages 770–81, March 2010. (Cit  en pages 15 et 38.)
- [Zhang 2008] Hui Zhang, Jason E Fritts et Sally A Goldman. *Image segmentation evaluation : A survey of unsupervised methods*. computer vision and image understanding, vol. 110, no. 2, pages 260–280, 2008. (Cit  en page 21.)

- [Zijdenbos 1994] Alex P Zijdenbos, Benoit M Dawant, Richard A Margolin et Andrew C Palmer. *Morphometric analysis of white matter lesions in MR images : method and validation*. Medical Imaging, IEEE Transactions on, vol. 13, no. 4, pages 716–724, 1994. (Cité en page 22.)
- [Zitová 2003] Barbara Zitová et Jan Flusser. *Image registration methods : a survey*. Image and Vision Computing, vol. 21, no. 11, pages 977 – 1000, 2003. (Cité en page 19.)

« HUIII »
