



HAL
open science

Caractérisation de la cocoteraie des Tuamotu à partir d'images satellites à très haute résolution spatiale.

Raimana Teina

► **To cite this version:**

Raimana Teina. Caractérisation de la cocoteraie des Tuamotu à partir d'images satellites à très haute résolution spatiale.. Interface homme-machine [cs.HC]. Université Pierre et Marie Curie - Paris VI, 2009. Français. NNT: . tel-00516952

HAL Id: tel-00516952

<https://theses.hal.science/tel-00516952>

Submitted on 13 Sep 2010

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

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

**THÈSE DE DOCTORAT DE
L'UNIVERSITÉ PIERRE ET MARIE CURIE**

SPÉCIALITÉ INFORMATIQUE

PRÉSENTÉE PAR

Raimana TEINA

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ PIERRE ET MARIE CURIE

Laboratoire d'accueil : Laboratoire d'Informatique de Paris 6 (LIP6)

Équipe d'accueil : PEQUAN- Département CALSCI

École Doctorale Informatique Télécom et Électronique de Paris

Titre de la thèse :

***Caractérisation de la cocoteraie des Tuamotu à partir d'images
satellites à très haute résolution spatiale.***

soutenue le 15 Septembre 2009 devant le jury composé de:

Jean-Luc	ZARADER	Président	
Jean-Paul	RUDANT	Rapporteur	UMLV/OTIG
Jocelyn	CHANUSSOT	Rapporteur	INPG/GIPSA-Lab
Jean-luc	ZARADER	Examineur	UPMC/ISIR Paris
René	ALT	Directeur de Thèse	UPMC/LIP6 Paris
Dominique	BÉRÉZIAT	Co-Directeur	UPMC/LIP6 Paris
Benoît	STOLL	Co-Directeur	UPF/GepaSud Tahiti

Remerciements

Je remercie Jean-Luc ZARADER, Professeur, qui me fait l'honneur de présider ce jury.

Je remercie Jean-Paul RUDANT, Professeur, et Jocelyn CHANUSSOT, Professeur, d'avoir bien voulu accepter la charge de rapporteur.

Je remercie Jean-luc ZARADER, Professeur, d'avoir bien voulu juger ce travail.

Je remercie René ALT, Professeur Emérite, qui a été mon directeur de thèse administratif.

Je remercie grandement Dominique Béréziat, co-encadrant au LIP6, pour tout l'encadrement qu'il a apporté pendant ma thèse, de m'avoir conseillé efficacement tout en me laissant libre cours à mes recherches et mes travaux, d'avoir été à l'écoute de mes attentes. Je le remercie aussi pour tout le soutien qu'il m'a apporté lors de ces travaux et surtout pendant certains moments difficiles.

Je remercie aussi Benoît Stoll, co-encadrant de l'Université de la Polynésie Française, pour toutes les missions et les encadrements apportés lors de mes séjours au sein du Laboratoire GePaSud.

Je remercie aussi tous les membres de l'équipe PEQUAN du LIP6 ainsi que ceux du département CALSCI pour leur soutien et les discussions enrichissantes que j'ai pu avoir avec eux.

Je remercie d'autre part Fred Jacq pour la mise à disposition de toutes ces données relevées sur le terrain. Je remercie aussi la population de Tikehau que nous avons dérangé pendant leurs travaux dans la cocoteraie lors de la mission sur le terrain, le maire de Tikehau qui a vu et qui était très intéressé par les premiers résultats sur le dénombrement des cocotiers. Je tiens aussi à remercier le Service du Développement Rural du Territoire (SDR) et le Service de l'Urbanisme pour toute l'aide apportée et pour la mise à disposition des images pour leur exploitation.

Je remercie aussi toutes les personnes dont j'ai omis le nom.

Dédicaces

*Je dédie cette thèse
à mes parents, Mareva et Alexandre, pour tous les sacrifices faits pour me donner les moyens d'arriver
jusqu'à la fin de cette thèse,
à ma chérie, Armelle, pour m'avoir encouragée jusqu'au bout, supportée lors de la rédaction et soutenue
tout le long de la thèse,
à mon grand-père, Jacques
et à mon ami Hervé.*

Table des matières

1	Introduction	5
1.1	Présentation du problème	5
1.2	Contributions	6
1.3	Plan du document	8
I	Problématique et données	9
2	Problématique	11
2.1	Contexte de l'étude	11
2.2	Problématique	14
3	Présentation des données des Tuamotu	15
3.1	Formation géologique des atolls de la Polynésie Française	15
3.1.1	Naissance des îles	15
3.1.2	Évolution des îles	16
3.1.3	Définition d'un atoll	16
3.2	Présentation des images satellites	16
3.2.1	Données de Tikehau	17
3.2.2	Données des autres îles	18
3.3	Adéquation données / problèmes	21
II	Dénombrement des cocotiers	23
4	État de l'art sur la détection des arbres	25
4.1	Recherche des maxima locaux du Laplacien	25
4.2	Recherche de motifs	26
4.3	Suivi de vallée	27
4.4	Algorithme de délimitation individuelle des arbres : TIDA	28
4.5	Ligne de partage des eaux	30
4.6	Processus ponctuels marqués	30
5	Dénombrement des cocotiers	33
5.1	Outils à la résolution du problème	33
5.1.1	Filtre Gamma	33
5.1.2	Analyse de texture : <i>Data Range</i>	34
5.1.3	Segmentation par ligne de partage des eaux	35
5.2	Délimitation des cocotiers en imagerie Ikonos	36
5.2.1	Architecture de la méthode	36
5.2.2	Masquage des zones végétales	37

5.2.3	Fusion des bandes par ACP	37
5.2.4	Lissage des images	38
5.2.5	Génération des marqueurs	38
5.2.6	Post-segmentation	39
5.2.7	Exploitation des résultats	40
6	Application à l’atoll de Tikehau et validation	43
6.1	Résultats et Validation	43
6.1.1	Résultats sur des zones de test	43
6.1.2	Validation par photo interprétation humaine	43
6.1.3	Validation par vérité terrain	44
6.1.4	Validation de la méthode sur les zones test du motu 16	46
6.1.5	Comptage exhaustif sur un <i>motu</i>	48
6.2	Traitement de l’ensemble des <i>motu</i>	50
6.2.1	Traitement du <i>motu</i> au Nord-Est (<i>motu</i> #13)	50
6.2.2	Traitement du <i>motu</i> principal (<i>motu</i> #16)	51
6.2.3	Résultats sur l’ensemble des <i>motu</i>	52
6.2.4	Cas des zones d’ombres induites par des nuages de basse altitude	52
6.3	Conclusion	55
7	Détection des ombres et estimation de la hauteur des arbres	57
7.1	Détection des ombres	57
7.1.1	La méthode de Polidorio	57
7.1.2	La méthode de Huang	59
7.2	Estimation de la hauteur à partir de l’ombre portée	60
7.2.1	Méthode d’estimation des hauteurs	60
7.3	Application à l’estimation des hauteurs des cocotiers	62
III	Typologie de la cocoteraie	65
8	État de l’art sur la classification	67
8.1	Définition de la classification	67
8.2	k -plus proches voisins	68
8.3	Classification par parallélépipèdes	68
8.4	Classification par distance minimum	69
8.5	Classification par distance de Mahalanobis	69
8.6	Classification Bayésienne	70
8.7	Séparateurs à Vaste Marge	71
8.7.1	Introduction	71
8.7.2	Le plan séparateur et la marge	72
8.7.3	Marge optimale dans le cas séparable	74
8.7.4	Erreur de classement minimal : cas linéaire non séparable	75
8.7.5	Séparation hautement non linéaire	76
8.7.6	Classification multi-classes avec les SVM	78
9	Choix du classifieur et des données d’entrée	79
9.1	Mesure de la séparabilité des classes	79
9.2	Choix des ensembles d’apprentissage	80
9.3	Influence des informations de texture sur la séparabilité des classes	82

10	Caractérisation des cocoteraies à partir de la texture	89
10.1	Codage de l'information	89
10.1.1	Ensemble d'apprentissage	89
10.1.2	Choix de l'information spectrale	90
10.2	Choix du classifieur	91
10.2.1	Analyse des performances	91
10.3	Application aux données Ikonos	92
10.3.1	Choix des paramètres du classifieur	92
10.3.2	Résultats	93
10.4	Conclusion	97
11	Caractérisation des cocoteraies par l'analyse de la distribution spatiale des cocotiers	99
11.1	Analyse spatiale des coordonnées	99
11.1.1	Processus ponctuel	99
11.1.2	Processus ponctuel de Poisson	99
11.1.3	Estimation de l'intensité par un SPPP	101
11.2	Application à la caractérisation de la cocoteraie	103
11.2.1	Estimation de la fonction d'intensité $\lambda(x)$	104
11.2.2	Analyse de second ordre de la cocoteraie	105
11.2.3	Classification des plantations de cocotiers	107
12	Conclusion et perspectives	111
IV	Annexes	115
A	Téledétection Optique	117
A.1	Le satellite Ikonos	117
A.2	Caractéristiques orbitales	117
A.3	Systèmes embarqués	118
B	Interaction avec la végétation	121
B.1	Caractéristiques spectrales de la végétation	121
B.1.1	Propriétés optiques des feuilles	121
B.2	Les différents indices de végétation	123
C	Génération des masques de végétation	125
D	Cartes de densité des secteurs test	139
E	Méthode utilisée pour l'interprétation visuelle	141
F	Correspondance des résultats avec les mesures sur le terrain	143
G	Cas des <i>motu</i> 13 et <i>motu</i> 16	147
G.1	<i>motu</i> 13	147
G.2	<i>motu</i> 16	148
H	Extraction des informations de texture	149
I	Estimation des surfaces de cocoteraies sur les atolls	153
J	Validation et extrapolation du nombres d'arbres sur Tikehau	157
J.1	Protocole des relevés	157
J.2	Interpolation des arbres	159

Bibliographie	166
Table des figures	167
Table des tableaux	171
Index	173
Résumé	173

Introduction

1.1 Présentation du problème

Dans ce manuscrit de thèse sont présentés les travaux effectués sur la caractérisation des cocoteraies (plantations de cocotiers) de l'archipel des Tuamotu, situé au Nord-Est de la Polynésie Française (voir la figure 1.1), à partir d'images satellites multi-spectrales Ikonos, sans bande proche infrarouge, à très haute résolution spatiale. Ce travail de recherche a contribué à l'élaboration d'un programme de régénération et de gestion des cocoteraies des atolls mis en place par le Service du Développement Rural (SDR) de la Polynésie Française.

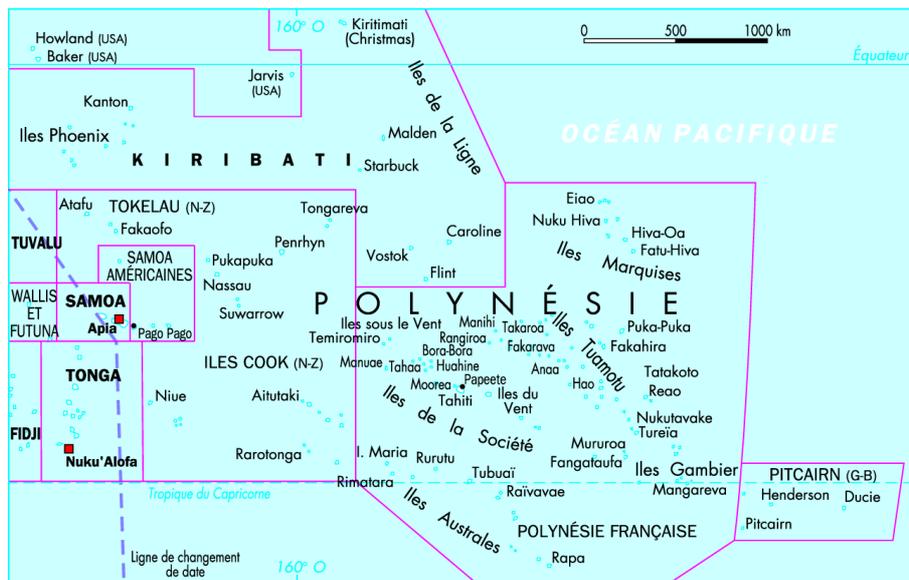


FIG. 1.1 – Carte de la Polynésie Française.

En dehors du tourisme, l'exploitation du *Coprah*, huile extraite à partir de la noix de coco, est l'une des principales ressources financières des atolls sur lesquels aucune activité de perliculture n'est recensée. Il est donc nécessaire, d'un point de vue économique, de réaliser un recensement complet des cocotiers afin d'estimer le volume d'huile pouvant être produit en fin d'année. En complément à ce recensement, l'estimation de la surface occupée par les plantations, ainsi que les types de plantations disponibles sur les atolls permettent d'obtenir un état des lieux de la situation servant de base pour le programme de renouvellement de la population de cocotiers. Des missions de recensement, sur le terrain, peuvent être effectuées mais cela demande du temps et des frais de mise en place importants. Grâce aux progrès effectués en imagerie satellite, il est maintenant possible d'obtenir des images pouvant aller jusqu'à 60 cm de résolution du pixel au sol, comme par exemple avec le satellite Quickbird. Quant à nous, nous disposons d'un jeu de données acquises par le satellite Ikonos offrant une résolution de 80

cm du pixel au sol, ce qui est suffisant pour résoudre le problème de recensement des cocotiers et de caractérisation de cocoteraies. L'utilisation de ces images, combinées à des techniques de segmentation adaptée, permet de faire un recensement des cocotiers et donc de réduire les coûts des missions de terrain, mais, bien entendu, la connaissance du terrain reste fondamentale pour valider les méthodes. Ce travail de recensement de la population de cocotiers sur les atolls des Tuamotu nécessite deux principales étapes :

- la segmentation individuelle des cocotiers afin d'estimer le nombre total d'arbres visibles sur les images,
- la classification des différents types de plantation en prenant en compte la densité locale, l'organisation spatiale des cocotiers.

Avant d'effectuer la segmentation des arbres, une étude préliminaire des images est nécessaire afin de localiser et classifier la végétation. Pour cela, il existe plusieurs indices, calculés à partir de l'information spectrale, qui caractérisent les types de végétation. Mais ces indices nécessitent la bande proche infrarouge pour pouvoir être calculés. Or, nous n'avons à notre disposition que des images prises dans le spectre visible, c'est-à-dire les canaux rouge, vert et bleu. L'absence de la bande proche infrarouge nous oblige alors à trouver des méthodes alternatives qui exploitent la structure spatiale visible dans l'image pour segmenter et classifier la végétation.

La littérature abonde en méthodes de segmentation d'arbres dans des images satellites. Certaines utilisent un modèle synthétique 3D d'arbre qui est projeté sur le plan de l'image en ayant correctement placé la source de lumière et le point de vue. Ce modèle génère un motif qui est alors recherché par corrélation dans l'image. Mais ce type de méthode présente l'inconvénient principal de ne détecter que le centre des arbres et non leur couronne. L'information apportée par la couronne est une source d'information importante sur le développement des palmes et l'état de santé de l'arbre. De plus, il est nécessaire de modéliser correctement l'arbre pour que la recherche par corrélation soit fructueuse. Une autre méthode, basée sur la recherche des maxima locaux du laplacien, permet de détecter des arbres en entier. Mais nous verrons que son application sur les images IKONOS de cocoteraies ne donne pas de résultats satisfaisants. Ceci est principalement dû au fait que les arbres ne sont pas suffisamment contrastés avec le sol (présence de sable fréquente sur le sol dans les atolls). D'autres méthodes utilisent l'hypothèse qu'il existe toujours des pixels plus sombres entre chaque arbre. En visualisant l'image comme un paysage montagneux, ces pixels sombres représentent les vallées. Un système de cheminement de ces vallées basé sur des règles et une stratégie de parcours dans le sens horaire permet de détecter les arbres et même de scinder en deux ceux qui se chevauchent. Sur le même principe, étant donné que les pixels sombres représentent les vallées, les plus clairs représentent une approximation des sommets des arbres. Ces maxima locaux, dans un voisinage de 7×7 pixels, servent de graines au processus de croissance de régions. Cette croissance s'arrête une fois un pixel d'une vallée rencontré ou lorsque ce pixel appartient déjà à une autre région. Une autre méthode, plus complexe, modélise l'image comme un processus ponctuel où chaque réalisation est un modèle d'arbre caractérisé par une ellipse et pour les régions où les arbres sont proches les uns des autres. Pour les arbres isolés, ce modèle est enrichi en prenant en compte son ombre portée au sol. Cette méthode n'est pas utilisée dans cette thèse car il est difficile de modéliser correctement la forme d'un cocotier, plus généralement les troncs ne sont pas rectilignes et peuvent présenter des configurations de croissance très différentes.

1.2 Contributions

La première contribution de cette thèse est une méthode de recensement des cocotiers ne nécessitant pas le canal proche infrarouge et basée sur l'hypothèse qu'il existe des pixels sombres entre les arbres. En effet, cette hypothèse est conforme à ce que l'on peut observer dans les images Ikonos. Nous construisons un réseau de tous les pixels qui appartiennent aux vallées et recherchons les maxima locaux dans une fenêtre de 7×7 pixels. Le réseau des minima ainsi que les maxima locaux servent alors de marqueurs, c'est-à-dire de source d'immersion, pour la segmentation par ligne de partage des eaux. Une fois la couronne de chaque arbre segmentée, nous calculons le centre de gravité pondéré par les intensités des pixels des trois canaux afin d'affiner l'approximation du centre. Pour chaque arbre, des informations statistiques peuvent être extraites telles que la couleur moyenne de la palme dans chaque

canaux, la surface mais aussi la position de chaque arbre et la densité locale des arbres. Afin de valider la méthode, la segmentation a été appliquée à l'image de l'atoll de Tikehau sur lequel une mission sur le terrain a été programmée. Des relevés de densité ont été réalisés et repérés au moyen d'un GPS. Les informations sur la topologie des arbres en fonction de leur voisinage a permis de mettre en évidence la présence d'un chevauchement important dans les cocoteraies artificielles et certains arbres se trouvent sous leurs voisins donc ils ne sont pas visible depuis le satellite. Les relevés sur le terrain sont comparés aux cartes de densité calculées à partir du résultat de la segmentation.

La seconde contribution de cette thèse concerne le traitement des zones ombragées par des nuages. En effet, l'ombre portée des nuages fait échouer le processus de classification de la végétation et donc biaise le résultat du dénombrement puisque ces zones ombragées sont pas traitées. Nous proposons une méthode très simple qui permet de segmenter correctement les cocotiers dans ces zones. Là encore, notre approche diffère de celles de la littérature par le fait qu'elle n'utilise pas la bande proche infrarouge.

Sous certaines conditions, il est possible d'utiliser l'ombre des cocotiers, facilement détectable, pour estimer leur hauteur. Ceci est possible si l'ombre portée au sol des cocotiers est visible dans l'image, la position du soleil est connue grâce aux méta-données disponibles dans les images. Sur ce principe, nous proposons une méthode d'estimation de la hauteur des cocotiers. Ce dernier point constitue une troisième contribution, mineure, de cette thèse.

Le dernier volet de cette thèse concerne la caractérisation de la répartition spatiale des cocoteraies. En particulier, il s'agit de reconnaître les cocoteraies artificielles ou sauvages ainsi que des configurations intermédiaires qui corresponde à des zones de cultures laissés à l'abandon. Pour réaliser cette caractérisation, deux paradigmes sont examinés.

Le premier repose sur la caractérisation des cocoteraies à partir de leur signature visuelle (structure spatiale et radiométrique) dans les images, c'est-à-dire leur texture. Le problème est alors traité par un formalisme de classification supervisée. Les paramètres de la classification sont calculés à partir d'ensembles d'apprentissage choisis par l'opérateur. L'information est encodée à l'aide d'une série de descripteurs de texture pertinents pour la caractérisation des structures végétales. Une étude de qualité quasi-exhaustive sur les huit classifieurs les plus utilisés de la littérature a été menée pour déterminer les performances de chacun. Ceci constitue une quatrième contribution de cette thèse.

Une telle méthodologie implique la classification de chaque pixel dans les images et induit un fort coût algorithmique (plusieurs heures de calcul sur de grandes images). Plutôt que d'analyser la texture, un second paradigme consiste à examiner la position spatiale de chaque cocotier et d'analyser leur répartition dans le plan. Ceci est rendu possible grâce à l'algorithme de dénombrement des cocotiers proposé précédemment. Cette approche est pertinente car le procédé de dénombrement est rapide (en comparaison à une classification pixel à pixel). Pour caractériser spatialement la répartition des cocotiers, nous avons choisit de tester statistiquement cette répartition avec celle engendrée par un processus ponctuel de Poisson. La densité des arbres est alors caractérisée par le paramètre d'intensité du processus poissonien. Ce paramètre est estimé localement et à différentes échelles pour bien rendre compte de la variabilité de la densité sur le terrain. Cette partie constitue la cinquième et dernière contribution de cette thèse.

Après caractérisation des différents types de plantation, les résultats sont croisés avec des relevés de terrain sur l'ensemble de l'atoll de Tikehau. Ceci permet d'obtenir, d'une part, une validation supplémentaire de la méthode de dénombrement des cocotiers par l'analyse des différentes densités locales et, d'autre part, de calculer des taux d'erreur de détection en fonction du type de la plantation. Grâce à ces erreurs relatives au type de la plantation, l'extrapolation du nombre d'arbres se fait au plus proche de la réalité en appliquant une correction adaptative selon le type de plantation traité.

1.3 Plan du document

Ce manuscrit est organisé en trois parties. Dans la première partie, il est présenté le mécanisme de formation des atolls de la Polynésie Française. Ceci permet de bien comprendre la structure géologique du terrain. Ensuite, les données sont présentées en détaillant celles de l'atoll de Tikehau et celles de l'atoll de Vahitahi. Enfin, nous concluons cette partie en présentant les contraintes liées aux données et la faisabilité de la résolution des problèmes considérées au tenant compte des résolutions spatiales des images.

La seconde partie s'articule autour du dénombrement individuel des arbres. Un état de l'art des méthodes utilisées pour la segmentation des couronnes des arbres est présenté. Cet état de l'art propose une liste non exhaustive de méthodes : utilisation d'un modèle synthétique d'arbres et recherche de motif dans l'image, recherche de minima locaux du laplacien, segmentation à base de règles de cheminement de vallée pour détourner un arbre, segmentation par ligne de partage des eaux contrainte par des marqueurs afin de limiter la sursegmentation, ou encore segmentation par processus ponctuels marqués utilisant un modèle d'arbre et de son ombre portée au sol. Nous présentons notre contribution à l'état de l'art du dénombrement d'arbres dans les images sans canal proche infrarouge. Les résultats obtenus et leur validation sont ensuite présentés.

La troisième partie traite de l'analyse de l'organisation spatiale des plantations de cocotiers afin de classer les plantations suivant leur densité, leur alignement ou encore leur régularité. Nous présentons un état de l'art des méthodes de classification supervisées telles que la classification par parallélépipèdes, la distance minimale, la distance de Mahalanobis, la classification par maximum de vraisemblance et les séparateurs à vaste marge (SVM). Nous montrons alors que la caractérisation des cocoteraies est possible en utilisant comme information la texture, décrite par des matrices de cooccurrence calculées sur l'image multispectrale. Une nouvelle étude est réalisée pour déterminer le classifieur le plus performant. Enfin, la caractérisation des cocoteraies par processus ponctuel de Poisson est présentée pour résoudre le problème du temps de traitement observé lors de la classification sur les textures.

Nous avons fait le choix de séparer le dénombrement individuel des cocotiers de la typologie des plantations dans deux parties différentes car ce sont deux problèmes distincts. En effet, le premier s'articule essentiellement autour de la segmentation des arbres alors que le second utilise les méthodes de classification pour caractériser les types de plantation. Dans le dénombrement des cocotiers, la classification n'est utilisée que pour l'étape de masquage des zones végétales et cela pour palier l'absence du canal proche infra rouge. Cette étape aurait très bien pu se faire par seuillage du NDVI si nous avions le canal proche infra rouge.

Le lecteur trouvera en annexe des compléments aux notions utilisés dans cette thèse, les résultats complets des différentes méthodes utilisées pour la segmentation des zones végétales, la segmentation des arbres avec la mise en correspondance des mesures faites sur le terrain.

Première partie

Présentation de la problématique et des données

Problématique

Dans ce chapitre, nous présentons le contexte de notre étude en mettant en avant l'intérêt économique que représente l'exploitation du *Coprah*. Ensuite, nous présentons la problématique que nous souhaitons résoudre tout au long de cette thèse.

2.1 Contexte de l'étude

L'exploitation des cocoteraies permet l'extraction de l'huile de *Coprah* à partir de la noix de coco mais aussi l'utilisation du bois à partir des cocotiers séniles (projet de scierie sur l'atoll de Tikehau). L'exploitation du *Coprah* constitue l'une des principales ressources financières des atolls habités sur lesquels aucune activité perlière n'est recensée. Le *Coprah* est l'albumen séché de la noix de coco. Les noix de coco sont récoltées une fois arrivées à maturité. Elles sont ouvertes et laissées à sécher un moment pour faciliter le détrocage (voir la figure 2.1) qui est effectué manuellement. La chair extraite est ensuite séchée au soleil, sur des présentoirs (voir la figure 2.2), jusqu'à la disparition totale de sa teneur en eau qui ne doit pas dépasser 6% pour obtenir du *Coprah*. Le *Coprah* séché est ensuite rassemblé dans des sacs qui sont pesés avant leur vente et l'acheminement maritime vers Papeete où se trouve l'Huilerie de Tahiti.



FIG. 2.1 – Séchage préliminaire des noix de coco ouvertes avant détrocage de l'albumen.

La culture du *Coprah* occupe traditionnellement une position privilégiée dans le paysage agricole et économique de la Polynésie Française. Cette culture est considérée par les pouvoirs publics comme un moyen essentiel de maintien des populations rurales dans les îles, notamment dans les atolls les plus éloignés où la perliculture n'est pas pratiquée. Ainsi, dans le cadre de l'aménagement du territoire, l'administration soutient financièrement la filière afin d'en assurer la pérennité et de garantir un revenu



FIG. 2.2 – Séchage de la chair de coco au soleil.

minimum aux producteurs de *Coprah*. Ce dispositif de soutien repose sur un prix de vente administré et des subventions accordées à l'Huilerie de Tahiti, seul organisme habilité à acheter la production aux cultivateurs.

Le prix acquité par l'huilerie varie selon la qualité du *Coprah* livré par le cultivateur, il est de 100 FCFP le kilogramme pour le *Coprah* de première qualité (prix fixé depuis le 1^{er} Octobre 2004) et de 55 FCFP le kilogramme pour le *Coprah* de deuxième qualité, sauf pour les Marquises où ce prix est fixé à 65 FCFP. Le tableau 2.1 présente la production selon la qualité et la valeur du *Coprah* pour l'année 2007. Le tableau 2.2 présente la quantité en poids net qui est entrée en silo selon la qualité à l'Huilerie de Tahiti pour l'année 2007.

Origine	Quantités récoltées en Kg et prix payé au producteur en FCFP/Kg						Valeur à la production en millions de FCFP
	1 ^{ère} qualité	Prix 1 ^{ère} qualité	2 ^{ème} qualité	Prix 2 ^{ème} qualité	Total	%	
Îles du Vent	306 674	100	88	55	306 762	3,23%	30,67
Îles Sous le Vent	1 197 420	100	7 238	55	1 204 658	12,67%	120,14
Îles Australes	110 070	100	3 494	55	113 564	1,19%	11,20
Îles Marquises	1 934 659	100	55 496	65	1 990 155	20,93%	197,07
Tuamotu-Gambier	5 730 389	100	161 681	55	5 892 070	61,97%	581,93
Polynésie Française	9 279 212		227 997		9 507 209	100,00%	941,01

TAB. 2.1 – Production par archipel selon la qualité et la valeur du *Coprah* en 2007. Source : Huilerie de Tahiti, Service du Développement Rural. Quantité récoltées : poids net au départ des îles.

Origine	Quantités entrées en silo en Kg			
	1 ^{ère} qualité	2 ^{ème} qualité	Total	% perte
Îles du Vent	302 674		302 674	1,33%
Îles Sous le Vent	1 177 980	7 221	1 185 201	1,62%
Îles Australes	107 985	3 361	111 346	1,95%
Îles Marquises	1 866 081	52 826	1 918 907	3,58%
Tuamotu-Gambier	5 383 112	154 029	5 537 141	6,02%
Polynésie Française	8 837 832	217 437	9 055 269	4,75%

TAB. 2.2 – Quantités entrées en silo en Kg pour l'année 2007. Source : Huilerie de Tahiti, Service du Développement Rural. Quantité récoltées : poids net en silo.

La figure 2.3 présente l'évolution de la production de *Coprah* récolté par archipel entre 1970 et 2007.

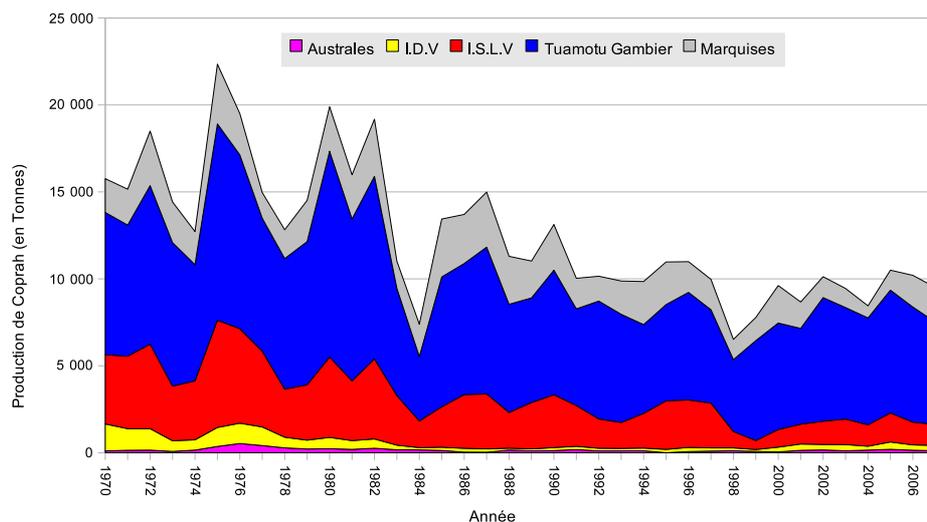


FIG. 2.3 – Évolution de la production d’huile de *Coprah* par archipel entre 1970 et 2007. Source : Huilerie de Tahiti, Service du Développement Rural.

L’Huilerie de Tahiti transforme le *Coprah* en huile brute et en tourteaux. L’huile brute est, pour l’essentiel exportée. Une faible part est raffinée sur place pour la fabrication du *Monoï de Tahiti*. Les tourteaux par contre servent de nourriture animale (bétail et volaille) et sont essentiellement destinés au marché local. Les utilisations de l’huile de coco sont principalement alimentaires : fabrication de margarines et graisses végétales (beurre de coco). L’huile de coco est aussi utilisée dans l’industrie de la savonnerie en raison notamment de sa teneur en acide laurique, qui donne au savon d’excellentes propriétés nourrissantes. L’huile de coco est aussi utilisé dans le secteur cosmétique de la filière *Coprah* en produisant le *Monoï de Tahiti* qui a reçu la distinction officielle de l’appellation d’origine en 1992. En Europe, il est couramment utilisé comme huile solaire hydratante et soin du corps.

La figure 2.4 présente l’évolution de la quantité d’huile de *Coprah* exporté (courbe en rouge) ainsi que la valeur qu’elle représente (courbe jaune). La production total est aussi affichée (courbe bleue) pour situer la part représentée par l’exportation. L’huilerie de Tahiti exporte principalement sa production vers l’Union Européenne. Le *Coprah* polynésien est en concurrence directe avec les productions des pays d’Asie du Sud-Est avec de gros producteurs comme les Philippines, la Thaïlande ou l’Inde.

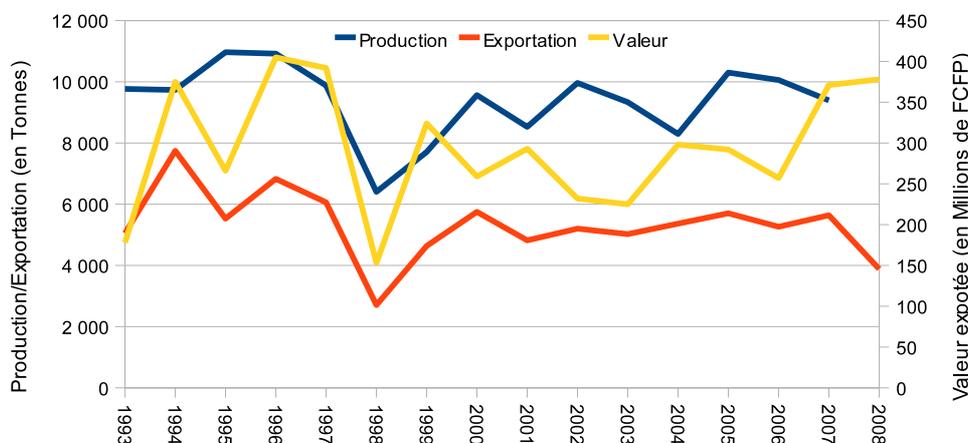


FIG. 2.4 – Évolution de l’exportation et de la valeur d’huile de *Coprah* entre 1993 et 2007. Source : Huilerie de Tahiti, Service du Développement Rural, Institut Statistique de la Polynésie Française.

Il est donc important pour l’économie des atolls, dont le *Coprah* est la principale ressource financière, d’avoir un recensement des cocotiers afin d’obtenir une estimation d’huile de *Coprah* pouvant être

potentiellement produite. La connaissance de l'état de chaque cocoteraie permet également d'optimiser la production.

Le protocole de Kyoto impose aux pays de faire un bilan carbone, donc des surfaces forestières de leur territoire. Une étude est en cours en Polynésie, mais à ce jour les cocoteraies ne sont pas comptabilisées pour éviter que des pays ne brûlent des forêts primaires et ne replantent des cultures de type cocoteraie, palmeraie à huile, etc. Cependant, les cocoteraies des Tuamotu devraient être considérées dans les surfaces de forêt naturelle, de part leur surface importante et la monotypicité du couvert végétal dans les atolls polynésiens, et le fait que les cocoteraies sont implantées depuis des siècles.

2.2 Problématique

La Polynésie Française se situe dans le Pacifique Sud à 18000 km de Paris (voir la figure 2.5). C'est un territoire composé de cinq archipels avec un total de 120 îles dont la plus importante et la plus peuplée est Tahiti. Les cinq archipels qui composent la Polynésie Française sont : l'Archipel de la Société (les Îles du Vent et les Îles Sous le Vent), l'Archipel des Marquises, l'Archipel des Australes et l'Archipel des Tuamotu-Gambier dont la plupart des îles sont devenues des atolls. Le territoire s'étend sur une surface de 2,5 millions de km², aussi grand que l'Europe sans la Russie, avec seulement 4000 km² de terre émergée, lagons compris. De part la configuration et l'éloignement des îles, l'organisation de mission sur les différents atolls est très coûteuse d'autant plus que certains atolls ne sont pas accessibles en avion mais uniquement par voix maritime.

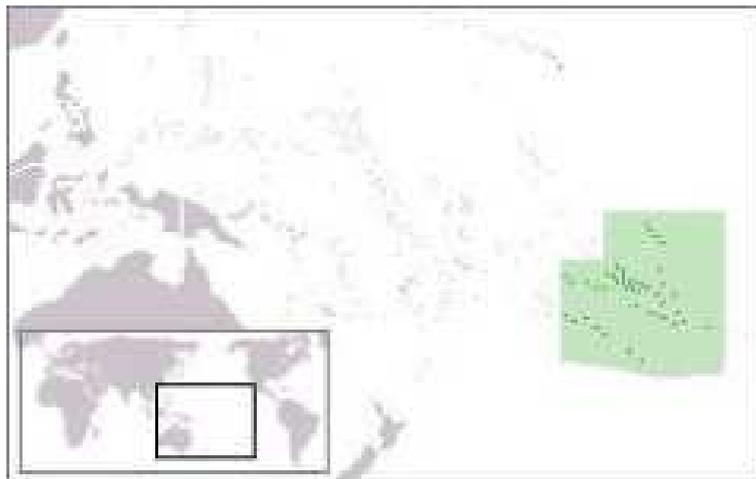


FIG. 2.5 – Carte de la Polynésie Française.

Grâce aux progrès réalisés dans le domaine de l'acquisition d'images satellites, il est possible d'obtenir des images d'une très grande résolution spatiale, telles que par exemple les images obtenues avec le satellite Ikonos dont la résolution maximale du pixel au sol est de 80 cm. Le Territoire de la Polynésie Française possède les droits d'exploitation sur de la donnée optique Ikonos de la quasi totalité des atolls de l'archipel des Tuamotu. À partir des informations contenues dans ces images, nous souhaitons résoudre les problèmes suivants :

- **La localisation** de chaque cocotier et l'identification de critères objectifs comme la taille du houppier, de la palme, la couleur de celle-ci et la densité locale de la cocoteraie,
- **la classification** des cocotiers et des cocoteraies selon leur organisation spatiale.

Afin de réaliser ces objectifs, nous devons exploiter au mieux l'information contenue dans les images satellites. Il est donc important de comprendre le processus physique qui résulte de l'acquisition et de la formation de ces images. Les chapitres suivants de cette partie détaillent ces processus.

Présentation des données des Tuamotu

Ce chapitre présente les données utilisées dans cette thèse ainsi que le contexte d'acquisition. Pour cela, il est important de comprendre la topologie du terrain et donc de la formation géologique des atolls. Nous présentons ensuite les données proprement dites.

3.1 Formation géologique des atolls de la Polynésie Française

3.1.1 Naissance des îles

Toutes les îles de la Polynésie Française sont d'origine volcanique. Ces îles reposent sur une grande plaque rigide de la lithosphère terrestre, la plaque Pacifique, engendrée à l'Est par la dorsale du même nom. Elle se déplace vers l'Ouest et le Nord-Ouest à raison d'environ 11 cm par an, et plonge sous la plaque Eurasienne, à l'Ouest, au niveau des grandes fosses du Pacifique (fosses des Aléoutiennes, des Mariannes, des Kouriles, etc).

Sur cette plaque, qui repose par plus de 4000 m de fond, naissent deux types de volcans :

- Ceux apparus au niveau de la dorsale du Pacifique-Est, zone de fracture favorable au passage de lave au travers de la lithosphère située au grand large de l'Amérique du Sud, puis qui ont dérivé. Il s'agit des plus vieux volcans de l'archipel, dont l'âge est très proche de celui de la plaque sur laquelle ils reposent (40 à 60 millions d'années). C'est le cas des îles de l'archipel des Tuamotu.
- Le deuxième type de volcan, dit de "point chaud", concernent toutes les autres îles de la Polynésie Française. Loin des limites des plaques, le magma des couches profondes remonte près du plancher océanique, le fait fondre, le perce, puis éjecte ses matériaux, parfois au-delà de la surface de l'océan, formant ainsi une île.

Le point chaud est fixe, le plancher océanique se déplaçant vers le Nord-Ouest, on assiste ainsi à la création d'un chapelet d'îles plus ou moins espacées en fonction du rythme des épisodes éruptifs. À titre d'exemple, l'archipel des Australes, situé au Sud de la Polynésie Française, a été créé par le point chaud de Mac Donald, toujours en activité et actuellement situé à 40 miles au Sud-Est de l'île de Rapa. Le sommet du volcan, actuellement en formation, n'est plus qu'à quelques dizaines de mètres de la surface de l'océan.

Les épisodes éruptifs qui ont donné naissance aux volcans polynésiens ont été brefs et considérables. Brefs, car la vitesse de déplacement de la plaque sur laquelle ils reposent étant relativement rapide, ces volcans ont été rapidement coupés de leur source de magma. On estime par exemple que le volcan principal de Tahiti a été créé en 0,75 million d'années seulement. Considérables, car ils ont éjecté en peu de temps des quantités formidables de lave. Le volcan de Tahiti a un volume estimé de 8 millions de m³. Sa hauteur totale à l'origine était de 12000 m. 3000 m pour la partie aérienne et 9000 m pour la partie sous-marine, si l'on tient compte de l'enfoncement que son poids a provoqué sur le plancher océanique.

3.1.2 Évolution des îles

À peine formé, le volcan aérien (ou île haute) est soumis à différents phénomènes qui vont progressivement le transformer en île basse ou atoll. C'est l'effet dit de "subsidence" causé à la fois par l'érosion aérienne, l'enfoncement de l'île, la formation de la barrière récifale et la dérive du plancher océanique. L'érosion aérienne est directement liée à la violence des phénomènes climatiques (pluies tropicales et vents principalement) qui lessivent ses flancs. L'enfoncement du volcan est dû à son propre poids ajouté à celui des coraux, qui poussent, faisant fléchir le plancher océanique sur lequel il repose. Tout au long de sa transformation, des coraux se développent en effet sur les côtes puis au large, formant une barrière et encerclant ainsi l'île d'un lagon. Une fois le volcan totalement effondré, il ne reste que l'anneau coralien qui émerge, appelé atoll, et qui disparaîtra lui aussi.

Le passage du stade d'île haute à celui d'atoll est rapide à l'échelle des temps géologiques (5 à 6 millions d'années). Aujourd'hui, on peut observer en Polynésie Française tous les intermédiaires entre ces types d'îles, y compris des configurations totalement immergées sous l'océan.

La figure 3.1 illustre le processus de formation d'un atoll.

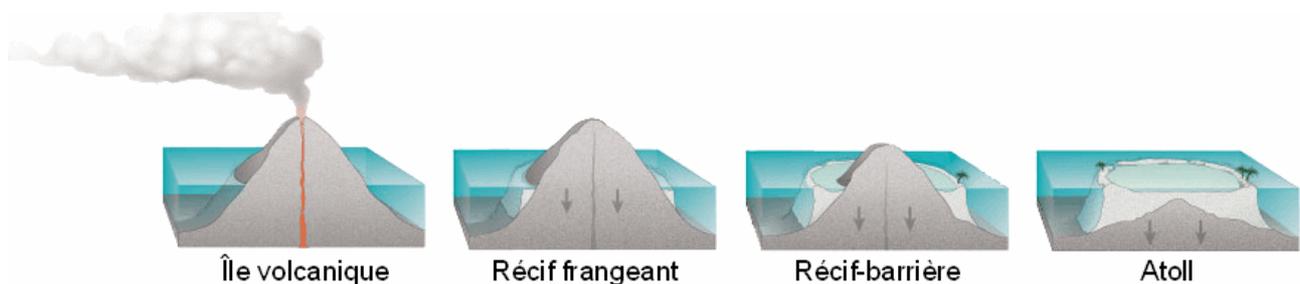


FIG. 3.1 – Schéma de la formation d'un atoll.[Field et al., 2002]

3.1.3 Définition d'un atoll

Un atoll (mot originaire des îles Maldives) est une île annulaire des mers tropicales constituée de récifs coralliens entourant une lagune centrale, le lagon. Plus scientifiquement, on parle de bioconstruction calcaire sur fond volcanique. Cette définition rend bien compte de la véritable nature d'un atoll. Des organismes minuscules, les coraux (ou plus exactement les madrépores) ont construit par empilement successif de leurs squelettes calcaires, des édifices atteignant des volumes colossaux. Les atolls sont aussi le résultat de la lutte que les coraux mènent pour rester à la lumière alors que leurs supports s'enfoncent inexorablement vers les profondeurs de l'océan. En Polynésie Française, les îlots qui composent un atoll sont appelés *motu*.

3.2 Présentation des images satellites

Pour notre étude, les données ont été récupérées auprès du Service de l'Urbanisme du Territoire de la Polynésie Française qui possède les droits d'exploitation de ces données. Les premières données mises à notre disposition sont celles de l'atoll de Tikehau. Puis d'autres données sont venues enrichir le catalogue pour d'autres traitements.

Toutes les données en notre possession ont été ortho-rectifiées. Ces images sont le produit d'une fusion entre les informations radiométriques des images multispectrales (à 3,2 m de résolution au nadir) et les informations spatiales de la bande panchromatique (à 82 cm de résolution au nadir) afin d'obtenir des images RVB (Rouge, Vert et Bleu) d'une résolution spatiale de 1 m ou 80 cm après rééchantillonnage. Cette fusion a été préalablement effectuée avant la réception des images. Nous précisons aussi que nous ne disposons pas du canal proche infrarouge.

Toutes les données en notre possession ont été acquises par le satellite Ikonos 2 dont les caractéristiques sont présentées en annexe A.

3.2.1 Données de Tikehau

Les premières données mises à notre disposition sont celles de l'atoll de Tikehau acquises par le satellite Ikonos 2 les 2 Juillet 2003 (à 20h35 GMT) et 20 Août 2003 (à 20h20 GMT). Ces données sont utilisées pour l'étude pilote de la cocoteraie des Tuamotu qui consiste à recenser et dénombrer les cocotiers présents sur l'atoll. Les images de Tikehau ont une résolution du pixel au sol de 1 m et sont géoréférencées dans le système WGS84, hémisphère sud et zone 6. La mosaïque complète de l'atoll Tikehau (figure 3.2) a une résolution de 28517 pixels par 28617 pixels et est centrée sur les coordonnées géographiques $148^{\circ}17'36.24''W$ et $14^{\circ}53'12.91''S$. De part la structure des atolls (présentée dans la section 3.1), il n'y a que 10% de données significatives car la plus grande partie de l'image est occupée par l'eau.

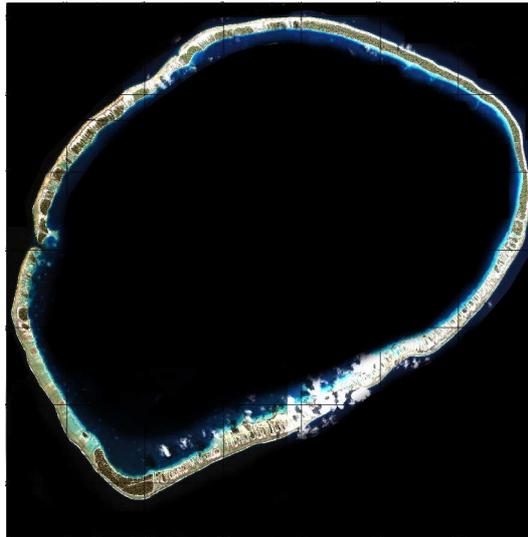


FIG. 3.2 – Vue globale de l'atoll de Tikehau.

Comme nous pouvons le constater sur la figure 3.2, l'eau occupe une grande partie de l'image. Nous avons donc fractionné l'image en zones comportant chacune un certain nombre de *motu*. Pour les identifier, nous leur avons attribué des numéros. La figure 3.3 montre les différentes zones marquées d'une lettre rouge et la numérotation des motus traités en jaune.

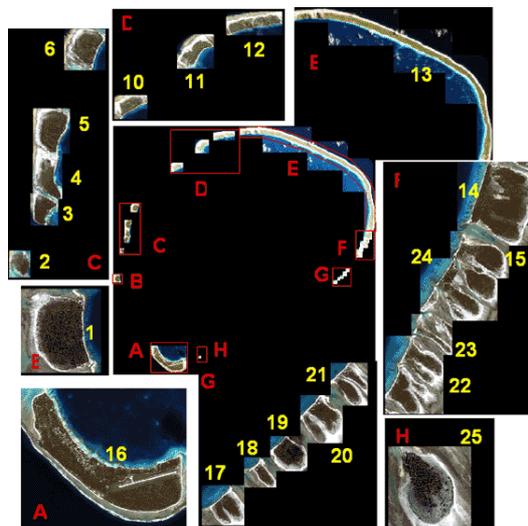


FIG. 3.3 – Fractionnement en zone de l'atoll de Tikehau : chaque zone (lettre en rouge) comporte un certain nombre de motu. Chaque motu est numéroté (chiffre en jaune).

La validation des résultats est réalisée sur des zones de test telles que celles définies sur la figure 3.4 grâce à une mission terrain. Ces zones de test sont représentatives des différentes configurations de plantations qu'il est possible d'avoir sur les images.



FIG. 3.4 – Motu principal : localisation des 6 zones de test du motu #16.

3.2.2 Données des autres îles

Panorama de ces données

Un autre jeu de données avec des images dont la résolution du pixel au sol est de 80 cm est également à notre disposition. L'acquisition des images a été réalisé entre Mars et Octobre 2005 en fonction de la douzaine d'atolls observés.

Ces données ne comportent cependant toujours pas de bande proche-infrarouge mais uniquement les canaux rouge, vert et bleu. Ces dernières sont utilisées pour l'estimation des surfaces de cocoteraies ainsi que la classification des types de plantation. Elles permettent également de montrer la robustesse des algorithmes proposés dans cette thèse. Elles illustrent aussi la pertinence de nos méthodes pour le traitement de grande quantité de données. Ces travaux rentrent dans le cadre d'un programme de régénération de la cocoteraie des Tuamotu conduit par le Service du Développement Rural (SDR) de la Polynésie Française. Le tableau 3.1 montre la liste avec les noms et date d'acquisition des images satellites et la figure 3.5 donne un panorama de ces données.

Atoll	Date	Résolution mosaïque	Coordonnées géographiques
Amanu	02/08/2005 19 :53 GMT	31308 × 29930 pixels	140°52'41.97''W,17°42'11.55''S
	26/10/2005 19 :49 GMT		
Fakahina	23/03/2005 19 :44 GMT	11077 × 6942 pixels	140°10'22.83''W,15°57'43.18''S
Fangatau	22/07/2005 19 :51 GMT	10135 × 6227 pixels	140°54'17.32''W,15°47'57.80''S
Nukutavake	05/06/2005 19 :40 GMT	7126 × 2689 pixels	138°48'47.55''W,19°16'14.48''S
Puka Puka	30/07/2005 19 :43 GMT	8352 × 7117 pixels	138°51'1.00''W,14°47'39.18''S
Reao	28/03/2005 19 :27 GMT	24142 × 16771 pixels	136°28'8.91''W,18°27'34.78''S
	02/06/2005 19 :31 GMT		
	05/07/2005 19 :33 GMT		
Tatakoto	05/06/2005 19 :40 GMT	17468 × 4761 pixels	138°27'32.39''W,17°19'36.04''S
Tureia	10/08/2005 19 :45 GMT	10473 × 16957 pixels	138°34'53.97''W,20°46'3.36''S
Vahitahi	01/09/2005 19 :46 GMT	10673 × 4120 pixels	138°51'52.96''W,18°45'35.19''S
Vairaatea	07/08/2005 19 :36 GMT	8983 × 6120 pixels	139°15'45.15''W,19°19'38.08''S
	09/09/2005 19 :38 GMT		

TAB. 3.1 – Liste des nouvelles données recueillies.

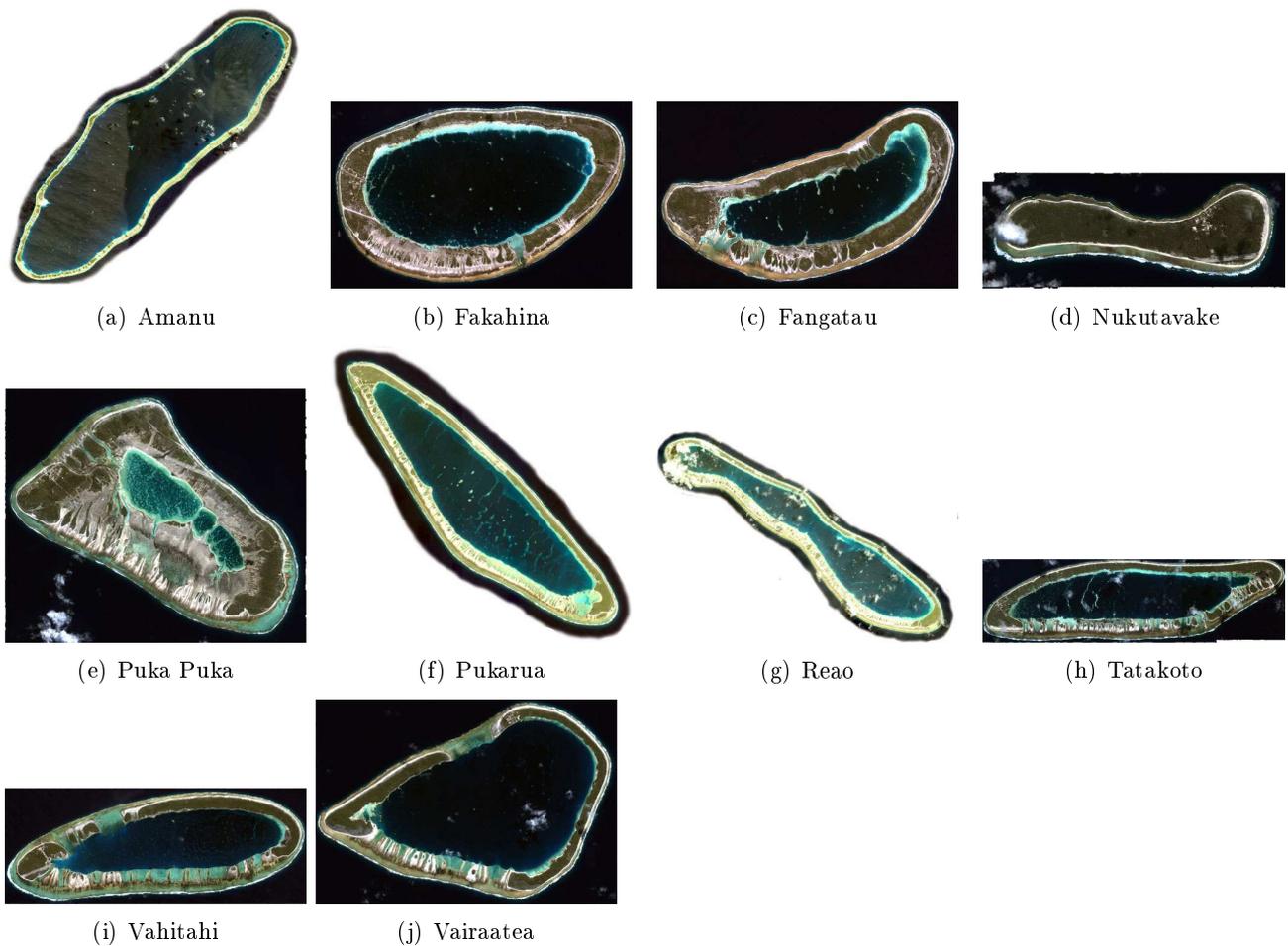


FIG. 3.5 – Panorama des autres données à notre disposition.

Présentation des données Ikonos de *Vahitahi*

Afin d'illustrer les méthodes utilisées pour résoudre le problème de classification des sols (voir le chapitre 9) et de caractérisation des types de cocoteraies (voir le chapitre 10), nous présentons en détail les données utilisées. La figure 3.6 montre une image Ikonos de l'atoll de Vahitahi situé à 1110 Km à l'Est de Tahiti dont les coordonnées géographiques sont $18^{\circ}47'$ de latitude Sud et $138^{\circ}50'$ de longitude Ouest. Le tableau 3.2 présente les statistiques extraites de l'image, telles que le minimum, le



FIG. 3.6 – Présentation de l'atoll de Vahitahi.

maximum, la moyenne et l'écart type de la distribution des niveaux de gris. Nous observons les canaux R, G et B sont fortement corrélés entre eux (tableau 3.2(b)). La figure 3.7 illustre l'histogramme associé. Le problème d'occupation des sols sera illustré sur une petite zone d'intérêt (voir la figure 3.8) pour

	Min	Max	Mean	StDev
R	1	2047	231,54	252,49
G	62	2047	390,17	267,82
B	116	2047	387,6	166,44

(a) Bases

	R	G	B
R	1		
G	0,95	1	
B	0,89	0,96	1

(b) Correlation

TAB. 3.2 – Statistiques de l'image 3.6.

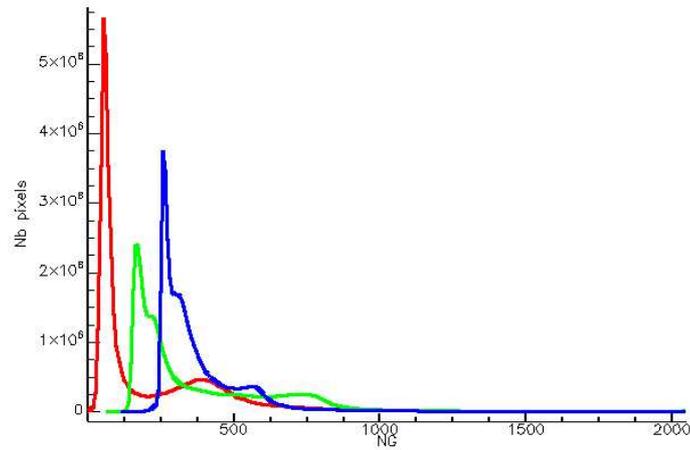


FIG. 3.7 – L'histogramme de l'image 3.6.

les expérimentations car les dimensions de l'image sont grandes et demandent beaucoup de ressources informatiques.



FIG. 3.8 – Visualisation d'une zone pour les illustrations futures.

La figure 3.10 ci-dessous présente la région du Nord-Est de l'atoll sur laquelle toutes les expérimentations concernant la caractérisation de la cocoteraie sera faite. Ce problème sera traité dans le chapitre 10.



FIG. 3.9 – Visualisation en détail de la zone du village.

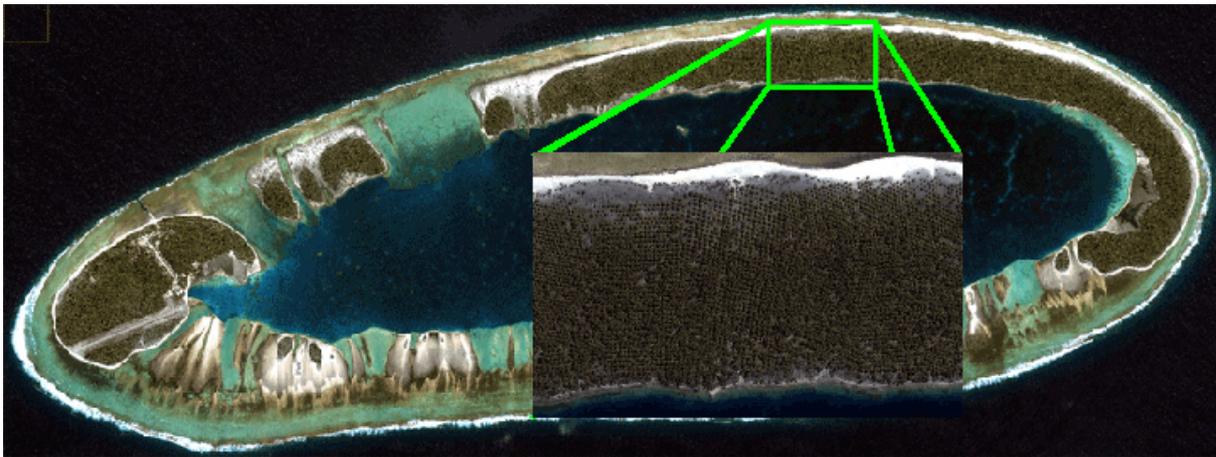


FIG. 3.10 – L'atoll de Vahitahi et un zoom de la zone d'étude au Nord-Est.

3.3 Adéquation données / problèmes

Lors de cette étude, nous nous sommes confrontés à plusieurs contraintes. Premièrement, nous n'avons aucune maîtrise sur la fusion des images multi-spectrales avec le canal panchromatique car celle-ci a été réalisée au préalable avant l'achat des images. Il en est de même pour le rééchantillonnage. Deuxièmement, les indices de végétation robusts (voir l'annexe B) sont calculés avec la bande proche infrarouge. Étant donné que nous ne disposons pas de cette donnée mais seulement des bandes spectrales rouge, verte et bleue, il nous faudra passer outre ces indices et trouver une approche alternative pour segmenter les régions végétalisées dans les images IKONOS. L'indice IP_{est} est le seul indice de végétation qui ne nécessite pas la bande proche infrarouge mais uniquement les canaux rouge et vert. Néanmoins, un tel indice n'est pas suffisant pour caractériser correctement les zones végétalisées.

La dernière génération des capteurs du satellite Ikonos permet, grâce à la grande résolution spatiale, d'observer des détails plus fins. Il est alors possible d'étudier et d'extraire de manière plus précise des objets. Le premier problème considéré est le dénombrement des cocotiers dans les images Ikonos. Ce dénombrement nécessite de segmenter individuel chaque arbre. Les palmes des cocotiers peuvent atteindre des longueurs de 6 m. En simplifiant la forme de la couronne des cocotiers par un cercle de rayon $2 \leq r \leq 6$ m, nous avons alors des surfaces de 12 m^2 à 113 m^2 à détecter. En prenant le cas des images de Tikehau, avec une résolution du pixel de 1 m, il faut alors détecter des régions de 12 à 113 pixels.

Deuxième partie

Dénombrement des cocotiers

État de l'art sur la détection des arbres

Ce chapitre présente un état de l'art non exhaustif des méthodes de segmentation de la végétation arborifère dans les images satellitaires. Cinq approche représentative sont présentées.

4.1 Recherche des maxima locaux du Laplacien

Dans [Karantzas and Argialas, 2004], les auteurs proposent de détecter les oliviers par la recherche des maxima locaux du laplacien. La méthode consiste dans un premier temps à réaliser une diffusion anisotropique [Perona and Malik, 1990]. L'image filtrée, I , vérifie l'équation aux dérivées partielles (4.1). Cette équation est discrétisée et on calcule ainsi l'image I au temps $t > 0$ à partir de l'équation (4.1) et de la condition initiale (4.2) qui est l'image originale I_0 . $g(\nabla I)$ est une fonction qui contrôle la diffusion et est choisie en fonction du gradient afin de préserver les contours, $\|\nabla I\|$ est la norme du laplacien de l'image et le coefficient κ contrôle la sensibilité sur les contours.

$$\begin{cases} \frac{\partial I(x, t)}{\partial t} = \operatorname{div} (g(\|\nabla I(x, t)\|) \nabla I(x, t)) & \forall x \in \Omega, \forall t > 0 & (4.1) \\ I(x, t_0) = I_0(x) & \forall x \in \Omega & (4.2) \end{cases}$$

avec

$$g(\nabla I) = \exp\left(-\left(\frac{\|\nabla I\|}{\kappa}\right)^2\right) \quad (4.3)$$

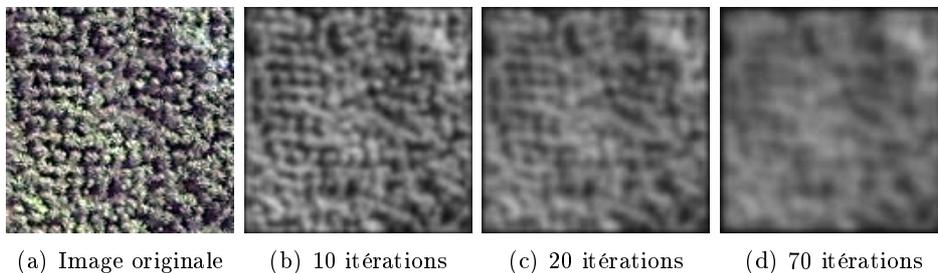


FIG. 4.1 – Exemple de filtrage par diffusion anisotrope : l'image originale et trois filtrage pour $t = 5, 10$ et 20 , avec la variable $\kappa = 25$.

L'image ainsi filtrée a pour propriété d'avoir ses contours renforcées et les régions homogènes lissées. Cet effet est d'autant plus net que le paramètre t est grand. La figure 4.1 montre les résultats d'un tel filtre pour différentes valeurs de t .

Dans un deuxième temps, une théorie connue sous le nom de détection de gouttes (*Blob Detection*) [Lindeberg, 1994] est utilisée afin d'extraire les arbres. Cette analyse repose sur le calcul du

laplacien de l'image. Soit I une image, on rappelle que la matrice hessienne est la matrice dont les éléments sont les dérivées partielles du second ordre et que le laplacien est la trace d'une hessienne (voir équation (4.4)).

$$H = \begin{pmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial I}{\partial x} \frac{\partial I}{\partial y} \\ \frac{\partial I}{\partial y} \frac{\partial I}{\partial x} & \frac{\partial^2 I}{\partial y^2} \end{pmatrix} \quad (4.4)$$

Un pixel dont le laplacien est nul et le déterminant de la matrice hessienne est positif est appelé point elliptique (à cause de la surface apparente de l'image dans cette région) et est concave ou convexe selon le signe du laplacien (voir la figure 4.2).

Extraire les gouttes dans une image revient à rechercher les extrema locaux du laplacien, ou de manière

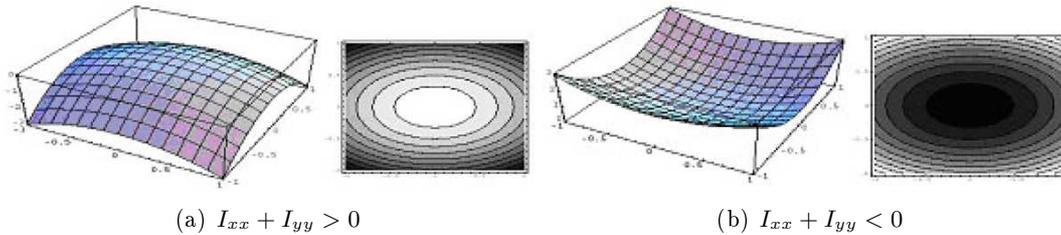


FIG. 4.2 – Apparence des points elliptiques en fonction du signe du laplacien [Karantzas and Argialas, 2004].

équivalente, à détecter les passages par zéro de son gradient tels que :

$$\nabla(\nabla^2 I) = 0 \quad (4.5)$$

Le signe de $\nabla^2 I$ détermine la clarté de la goutte : clair si $\nabla^2 I > 0$ et foncée si $\nabla^2 I < 0$

Cette méthode a été appliquée à la détection des oliviers sur des différentes images satellites à différentes résolutions du pixel au sol (1 m pour Quickbird, 4 m pour Ikonos et 10 m pour SPOT HSV) [Karantzas and Argialas, 2004]. Les étapes sont illustrées dans la figure 4.3. Or, il s'avère que dans ces images les houppiers des oliviers sont relativement bien contrastés par rapport au sol qui, lui, est très homogène, ce qui n'est pas le cas pour des images Ikonos de cocoteraies. Des tests ont été fait mais sans résultat probant.

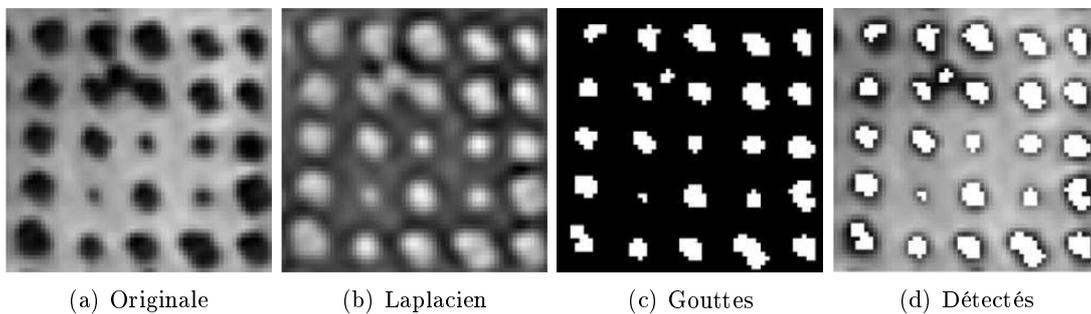


FIG. 4.3 – Application à l'extraction des oliviers d'après [Karantzas and Argialas, 2004].

4.2 Recherche de motifs

Les méthodes de recherche de motifs, développées initialement par [Pollock, 1996], reprises et améliorées par [Larsen, 1998, Olofsson, 2002, Olofsson et al., 2006], essaient de trouver dans l'image la meilleure correspondance entre un modèle synthétique d'arbre et une région dans l'image.

Dans [Pollock, 1996], il est montré que la couronne d'un cocotier est correctement approximé par une révolution autour d'un axe d'une courbe de type demi-ellipse modifiée dont l'équation est :

$$\frac{|z|^n}{a^n} + \frac{(|x|^2 + |y|^2)^{n/2}}{b^n} = 1 \quad (4.6)$$

$a, b \in \mathbb{R}_*^+$ et n un entier. Le paramètre n contrôle la forme générale et vaut typiquement : 1 (pour un cône), 2 ou 3 (pour une ellipsoïde). En changeant ces paramètres, il est possible de générer différents modèles en 3D pour différentes tailles de houppier. La figure 4.4 montre différents modèles générés ainsi que ces mêmes modèles illuminés par une lumière et projetés sur le plan de l'image. La première ligne a été générée en utilisant $a = 100$ et la seconde avec $a = 200$, dans tous ces exemples $b = 50$. Pour chacune des colonnes, n prend les valeurs 1, 2 et 3.

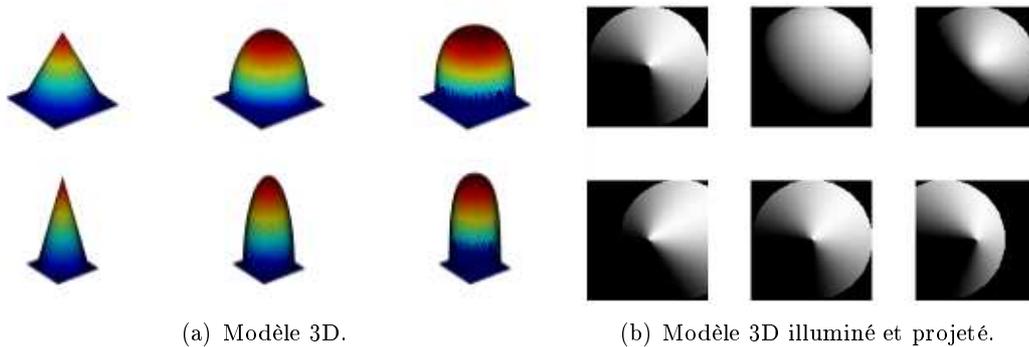


FIG. 4.4 – Modèle synthétique d'arbres [Pollock, 1996].

Pour déterminer la position des arbres, il est nécessaire de réaliser une recherche par maximum de corrélation de ces motifs. L'inconvénient de cette méthode est que le résultat indique uniquement les positions des arbres mais ne renseigne en aucun cas sur la forme de l'arbre détecté. De plus, le motif à rechercher dans l'image varie en fonction de la taille et la forme des arbres à détecter ce qui implique d'avoir autant de modèles que de types d'arbres à détecter et nécessite de fait une étape de calibration.

4.3 Suivi de vallée

La technique de suivi de vallée [Gougeon, 1995, Gougeon et al., 2003] considère les niveaux de gris comme des altitudes. De cette façon, les contours sombres bordant chaque arbre sont des minima locaux et autant de vallées. La méthode réalise quatre parcours de l'image : haut-gauche, vers bas-droite, bas-droite vers haut-gauche, bas-gauche vers haut-droite et haut-droite vers bas-gauche en recherchant si, par rapport à son voisinage, le pixel courant n'est pas un point d'une vallée, c'est-à-dire un pixel dont le niveau de gris est un minimum local dans ce voisinage. La figure 4.5 présente les différents voisinages utilisés en fonction du sens du parcours. Le pixel bleu est le pixel en cours d'analyse et en rouge ce sont les pixels pris dans le voisinage pour comparer les valeurs de niveaux de gris. La direction du parcours de l'image indiquée par la flèche.

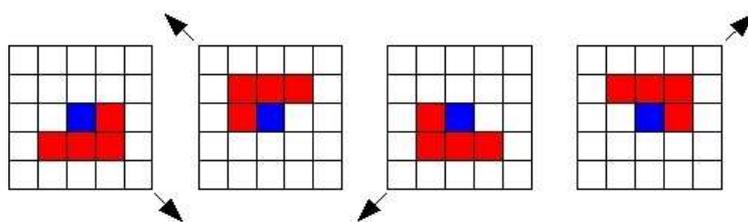


FIG. 4.5 – Voisinage utilisé lors des passages dans les 4 directions

Après ces quatre parcours, on obtient un réseau minimal de vallées sous forme d'image binaire. On

applique ensuite un ensemble de règles pour élaguer ce réseau et obtenir des couronnes. Ces règles, au nombre de cinq, définissent les sens de rotation et d'avancée possible le long du contour d'une région :

- Règle 1 : assure que le suivi du contour se fait dans le sens horaire et sans discontinuité,
- Règle 2 : autorise un écart de 45° dans le sens anti-horaire dans certains cas,
- Règle 3 : gère les virages à 90° dans le sens anti-horaire en vérifiant que le pixel fait partie du contour ou si celui-ci est un trou,
- Règle 4 : gère les tours à 135° qui peuvent être soit des anomalies soit une séparation du contour en deux qui caractérise deux arbres,
- Règle 5 : gère les retours en arrière qui signifie généralement la présence de deux arbres.

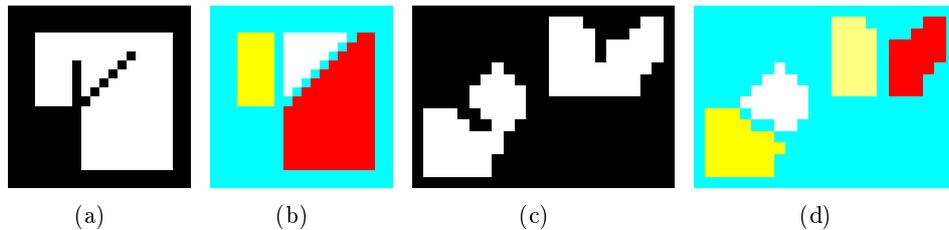


FIG. 4.6 – Application des règles de suivi de contour sur des images synthétiques.

Bien que les règles aient été validées sur des images synthétiques (voir la figure 4.6) et que nous ayons suréchantillonné nos images à une résolution du pixel au sol de 32 cm de manière à avoir la même résolution que celle utilisée par l'auteur, la méthode n'a pas donné de résultats probants sur les images Ikonos (voir la figure 4.7). En particulier, la méthode échoue lorsque les cocotiers sont trop agglomérés.

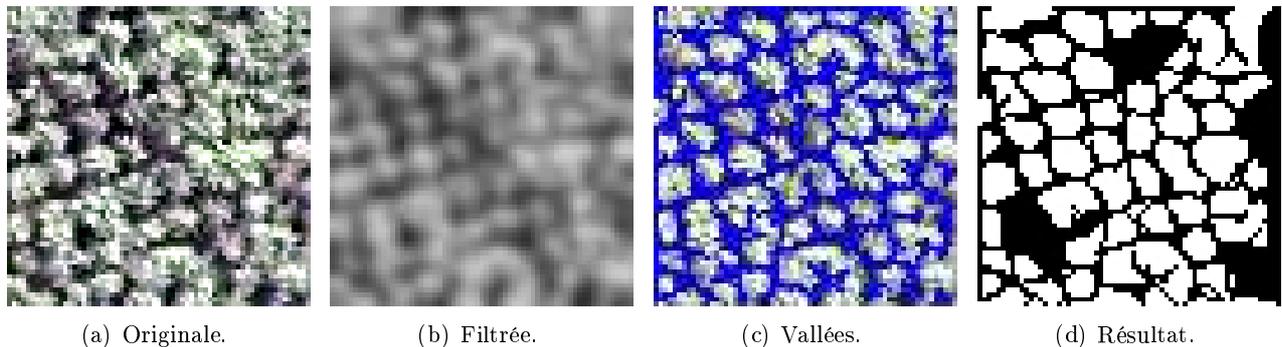


FIG. 4.7 – Application de la méthode sur une image réelle.

4.4 Algorithme de délimitation individuelle des arbres : TIDA

La méthode TIDA [Culvenor, 2002] construit un réseau de points de contour qui correspond aux couronnes des arbres. Les points de contours sont les minima locaux calculés dans un voisinage de 7×7 pixels et dans quatre directions : verticale, horizontale et les deux diagonales (voir la figure 4.8). Des maxima locaux sont recherchés suivant le même principe et ces derniers servent par la suite de graines pour le processus final.

À chaque maximum local on associe une fréquence $f_{LocalMax}$ d'identification qui indique le nombre de fois où ils sont identifiés comme maximum local pour les quatre directions de recherche. La figure 4.9 montre des images de maxima locaux pour différentes fréquences. Les pixels dont la fréquence $f_{LocalMax}$ est supérieure ou égale à 3 sont de bons candidats pour être des centres d'arbre.

D'autre part, la construction d'un réseau de minima est réalisée en utilisant les quatre parcours présentés précédemment. On retire du réseau les pixels qui ont le moins de chance d'appartenir au contour d'un arbre. Ce processus d'ajustement s'effectue en trois étapes :

1. fermer les contours à l'aide de règles sur le voisinage,

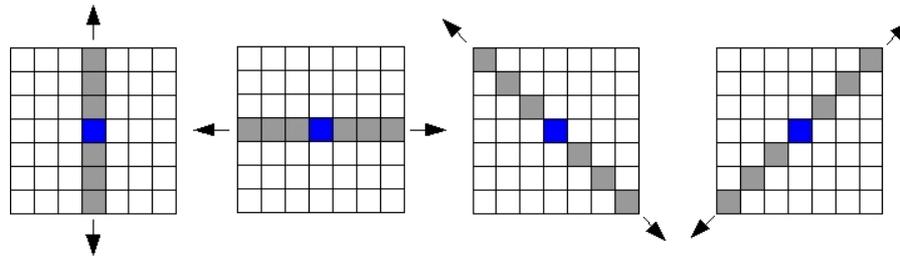


FIG. 4.8 – Les 4 directions définissant le domaine de recherche d’un maximum local.

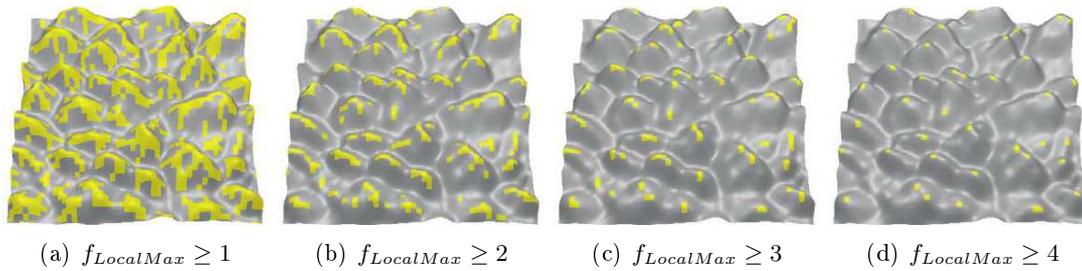


FIG. 4.9 – Identification des maxima locaux en fonction de leur fréquence d’identification. Les pixels en jaune sont ceux identifiés comme des maxima locaux [Culvenor, 2002].

2. retirer les pixels isolés qui n’appartiennent pas au réseau des minima,
3. affiner le contour de sorte qu’il fasse un pixel d’épaisseur.

Un pixel n’appartenant pas à un contour est reclassifié comme appartenant au contour si au moins deux de ces voisins appartiennent à ce contour et que ces voisins ne sont pas contigus. La figure 4.10 présente les règles utilisées pour la reclassification d’un pixel.

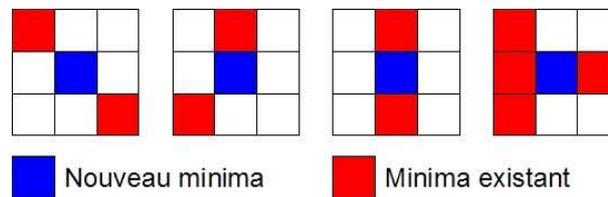


FIG. 4.10 – Exemple de cas de reclassification d’un pixel comme appartenant au réseau des minima.

Ensuite, il faut procéder à l’élimination des pixels isolés considérés comme des minima mais n’appartenant pas au réseau. Ceci est réalisé en retirant les pixels qui ont moins de deux pixels immédiats qui ne font pas partis du réseau. La figure 4.11 présente le résultat de l’ajustement réalisé.

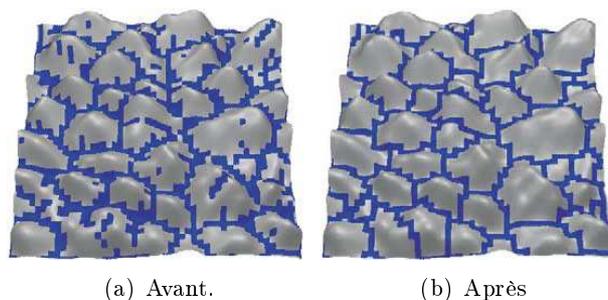


FIG. 4.11 – Avant et après le processus d’ajustement [Culvenor, 2002].

L’étape finale est la détection des couronnes des arbres par croissance de région à partir des maxima locaux détectés avec les critères d’arrêt suivants :

- rencontre avec le réseau de minima,
- rencontre d'un pixel dont le niveau de gris est en dessous de la moyenne des centres des arbres à un facteur multiplicatif près fixé empiriquement (voir la figure 4.12),
- rencontre avec une autre région.

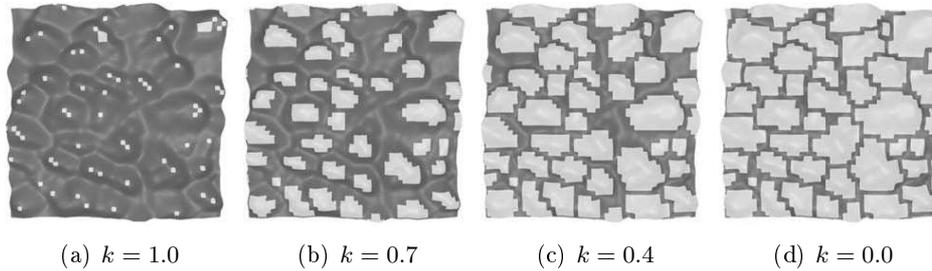


FIG. 4.12 – Effet de la constante min_k sur le résultat de la croissance de région [Culvenor, 2002].

4.5 Ligne de partage des eaux

La ligne de partage des eaux (LPE) [Beucher and Lantuéjoul, 1979, Vincent and Soille, 1991] est très utilisée pour segmenter des arbres dans des images satellitaires. Citons [Komura et al., 2002, Komura et al., 2003, Komura et al., 2004] dans lesquelles les auteurs proposent de représenter la couronne d'un arbre par un cercle. Ils utilisent ensuite la LPE pour séparer les régions circulaires segmentant ainsi les arbres. Dans [Wang et al., 2004], les auteurs proposent d'utiliser la LPE contrainte par des marqueurs pour segmenter les arbres.

Nous allons utiliser une variante de ces méthodes que nous présenterons au chapitre 5.

4.6 Processus ponctuels marqués

Dans [Perrin et al., 2006], les auteurs proposent d'extraire les houppiers des arbres en modélisant les peuplements forestiers par un processus ponctuel [Baddeley and van Lieshout, 1993, van Lieshout, 2000] marqué d'ellipses. Les marques d'une ellipse sont sa position (qui est aussi celle de l'arbre), son petit et son grand rayon ainsi que son orientation (voir la figure 4.13).

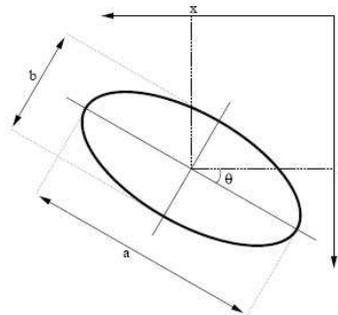


FIG. 4.13 – Position et marques d'une ellipse [Perrin et al., 2006].

Un arbre est modélisé en considérant uniquement leur ombre. Ainsi, un arbre est considéré comme une forme claire entourée d'une zone d'ombre (voir la figure 4.14).

Les positions des arbres sont générés au moyen de processus ponctuels (voir le chapitre 11 pour une introduction aux processus ponctuels) et adapté en utilisant deux types de mouvements sont considérés :

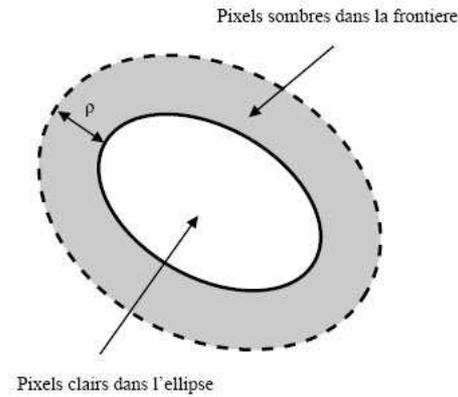


FIG. 4.14 – Modélisation d'un arbre avec son ombre [Perrin et al., 2006].

- ceux qui ne changent pas le nombre d'arbres détectés : ils modifient uniquement la position ou les marques des objets. Ces mouvements sont la rotation (modification de l'orientation), la translation (modification de la position), la dilatation (modification du petit axe et du grand axe) ainsi que toute combinaison de ces mouvements (voir la figure 4.15).

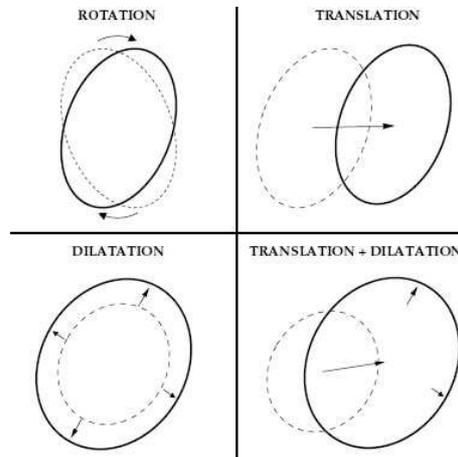


FIG. 4.15 – Exemples de mouvements simples [Perrin et al., 2006].

- ceux qui changent le nombre d'arbres détectés : ces types de mouvements permettent de fusionner ou diviser des objets afin d'éviter une sur-détection ou une sous-détection (voir la figure 4.16).

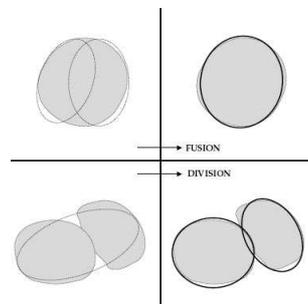


FIG. 4.16 – Exemple de fusion et de division [Perrin et al., 2006].

Les résultats obtenus en utilisant les mécanismes décrits ci-dessus sont présentés dans la figure 4.17.

La gestion des arbres isolés utilise une modélisation particulière : en effet, un arbre isolé se distingue uniquement par son ombre et non plus par une couronne ombragée qui l'entoure (voir la figure 4.18). Un exemple de résultat de la détection d'arbres isolés est présenté dans la figure 4.19.

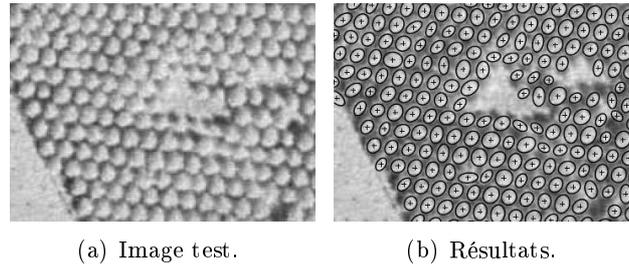


FIG. 4.17 – Exemple de résultats obtenus sur des données satellites.

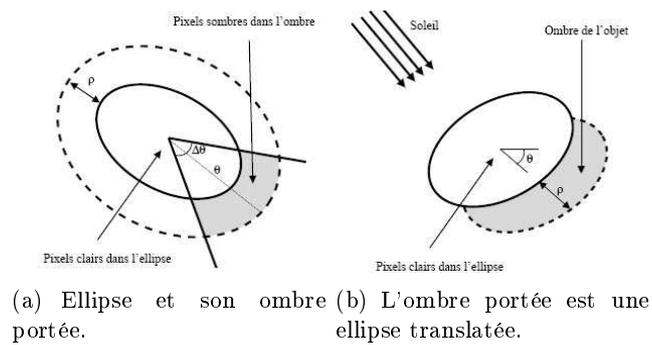


FIG. 4.18 – Modèle pour un arbre isolé.

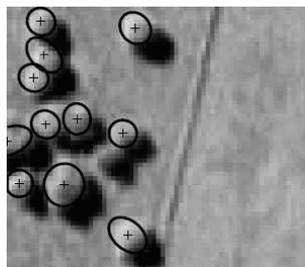


FIG. 4.19 – Résultat de la détection d'arbres isolés.

Dénombrement des cocotiers

Ce chapitre présente notre contribution à l'état de l'art pour la segmentation des cocotiers. Cette contribution se distingue par le fait qu'elle n'utilise pas (par force de nécessité) la bande PIR mais plutôt l'information spatiale structurelle des images. De ce fait, l'algorithme se structure en quatre étapes principales qui sont : pré-traitement et masquage des zones végétales, détection des arbres, segmentation des cocotiers et post-traitement. Certaines de ces étapes reposent sur des techniques de l'état de l'art que nous présentons dans une première section. L'architecture de la méthode est précisée dans la seconde section et l'exploitation des résultats est montré dans la troisième section.

5.1 Outils à la résolution du problème

Il est présenté dans cette section les outils utilisés pour résoudre notre problème de détection des cocotiers. Ces principaux outils sont : le filtre gamma, l'analyse de texture et la ligne de partage des eaux pour segmenter les cocotiers.

5.1.1 Filtre Gamma

Le filtre Gamma réalise un filtrage Gamma par Maximum A Posteriori (MAP). Le filtrage Gamma est principalement utilisé sur des données radar, donc en niveaux de gris, afin d'atténuer le bruit des hautes fréquences (*Speckle*) tout en préservant les éléments dans les hautes fréquences (c'est-à-dire les contours).

Le filtre Gamma MAP a été proposé par [Kuan et al., 1985]. L'application d'un tel filtre nécessite une connaissance *a priori* des fonctions de densité de probabilité de la distribution des niveaux de gris de l'image. Cette distribution des niveaux de gris est supposée suivre une distribution gaussienne. Cependant, ce n'est pas tout à fait réaliste puisque cette hypothèse suppose implicitement des niveaux de gris négatifs. Les auteurs de [Lopes et al., 1990] proposent de modifier le filtre de Kuan en supposant cette fois-ci que la distribution des niveaux de gris dans l'image suit une loi Gamma et mettant en place deux niveaux de seuillage. La définition d'un filtre Gamma MAP est donnée par l'équation (5.1) :

$$R = \begin{cases} \bar{I} & \text{for } C_I \leq C_u \\ \frac{(\alpha-L-1)\bar{I} + \sqrt{\bar{I}^2(\alpha-L-1)^2 + 4\alpha L\bar{I}}}{2\alpha} & \text{for } C_u \leq C_i \leq C_{max} \\ I_{cp} & \text{for } C_i > C_{max} \end{cases} \quad (5.1)$$

avec : $L = NLOOK$ le nombre d'observations de la même scène (pour des images d'intensités de niveaux de gris, $L = 1$), $C_{max} = \sqrt{2}C_u$ et $\alpha = \frac{1+C_u^2}{C_I^2-C_u^2}$, $C_u = \frac{\sigma_u}{\bar{u}}$ et $C_I = \frac{\sigma_i}{C_i}$. \bar{u} est la moyenne des niveaux de gris de la fenêtre glissante utilisée pour le filtrage avec une variance σ_u^2 . \bar{I} est la moyenne des niveaux de gris de l'image et C_{cp} est la valeur de niveaux de gris du pixel centré de la fenêtre de filtrage.

Le taille de la fenêtre du filtre Gamma varie entre 3 et 11. Si le filtre est trop petit, l'algorithme de filtrage ne sera pas efficace. Par contre si le filtre est trop grand, des détails de l'image seront perdus

lors du processus de filtrage. Un filtre de taille 7×7 donne généralement de bons résultats.

Ce filtre minimise la perte d'information de texture sur une image dont la distribution statistique suit une loi Gamma mieux que ne le font les filtres de Frost [Frost et al., 1982] ou de Lee [Lee, 1986]. Ce filtre convient à un grand nombre d'applications dont la distribution statistique suit une loi Gamma tel que dans l'industrie forestière, en agriculture et pour les applications aux océans. Le filtre conserve la valeur des pixels lorsque la distribution statistique ne suit pas la loi Gamma.

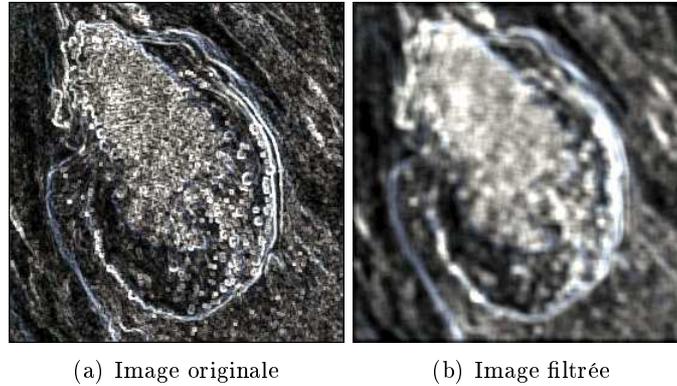


FIG. 5.1 – Exemple d'utilisation du filtre Gamma : à gauche l'image originale (résultat d'un filtre de texture : le *Data Range*) et à droite l'image filtrée avec filtre Gamma en utilisant une fenêtre de 7×7 pixels.

5.1.2 Analyse de texture : *Data Range*

La classification de la mer, le corail et la végétation peut être réalisée facilement grâce à une classification bayésienne (voir la section 8.6) avec apprentissage. Concernant la végétation, seuls les arbres nous intéressent. Pour masquer les zones de végétation non significatives dans les images, nous appliquons un filtre de texture appelé *Data Range* [Anys et al., 1994] sur une fenêtre de 3×3 pixels (voir la figure 5.2 a)). Le *Data Range*, calculé à partir des matrices d'occurrence, permet de mettre en évidence des zones à forte variance locale. Ensuite nous filtrons cette image avec un filtre Gamma avec une fenêtre de 7×7 pixels. Un profil de la zone avec de la végétation haute et de la végétation basse nous permet de visualiser la valeur de seuil qui nous permettra de créer un masque contenant uniquement de la végétation haute. Le tableau C.1 de l'annexe C indique les différentes valeurs de seuillage utilisées pour chacun des *motu*.

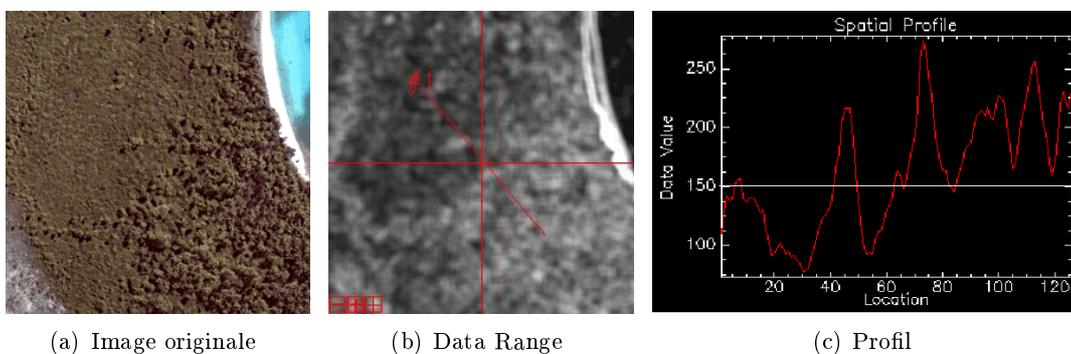


FIG. 5.2 – Exemple de segmentation de texture : l'image originale, l'image obtenue avec un filtre de texture *Data Range* avec la position d'un transec et le profil du transec pour visualiser la valeur de seuillage.

5.1.3 Segmentation par ligne de partage des eaux

La ligne de partage des eaux (LPE) considère les images en termes géographiques. Une image peut en effet être perçue comme un relief si l'on associe le niveau de gris de chaque pixel à une altitude. La ligne de partage des eaux est donc la crête formant la limite entre deux bassins versants (voir la figure 5.3). Pour l'obtenir, il faut imaginer l'immersion d'un relief dans de l'eau, en précisant que l'eau

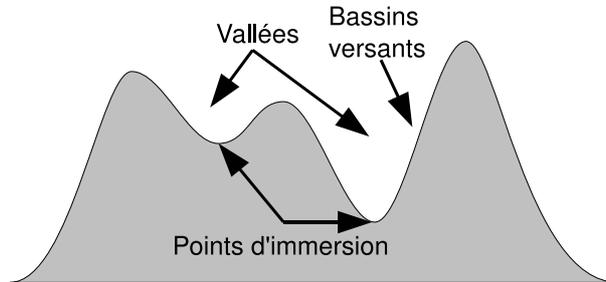


FIG. 5.3 – Terminologie utilisée pour la LPE.

ne peut pénétrer dans les vallées que par ses points d'immersion. La ligne de partage des eaux est obtenue par les points où deux lacs disjoints se rejoignent au cours de l'immersion (voir la figure 5.4). Après immersion, on obtient des régions homogènes en niveau de gris.

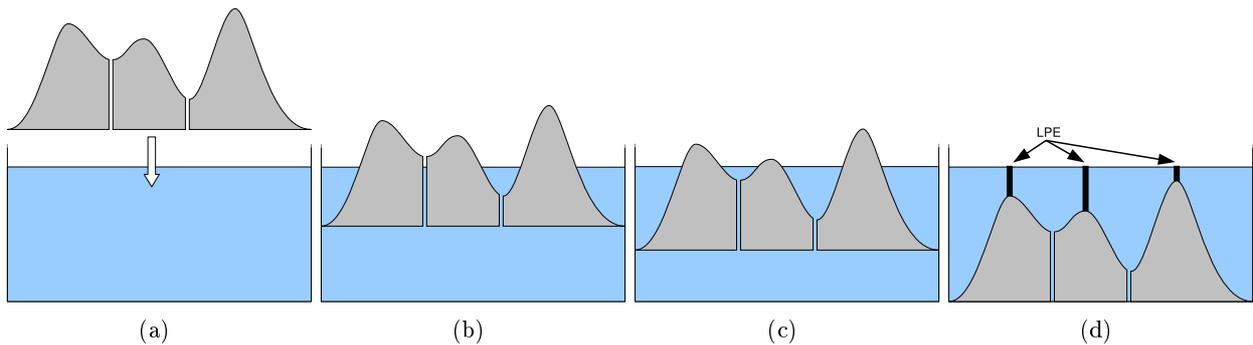


FIG. 5.4 – Processus d'immersion de la LPE.

Dans la méthode historique [Vincent and Soille, 1991], les points d'immersions sont donnés par les minima locaux. Dans la pratique, un tel choix conduit à une forte sur-segmentation puisqu'à chaque minimum local correspondra une région. Pour limiter cette sur-segmentation, il suffit de limiter le nombre de points d'immersion. On parle de "contraindre" la LPE. La figure 5.5 montre ainsi comment fusionner deux régions, en ne choisissant qu'un seul des deux minima locaux. Par la suite, on parlera de "marqueurs" les points d'immersion effectivement retenus.

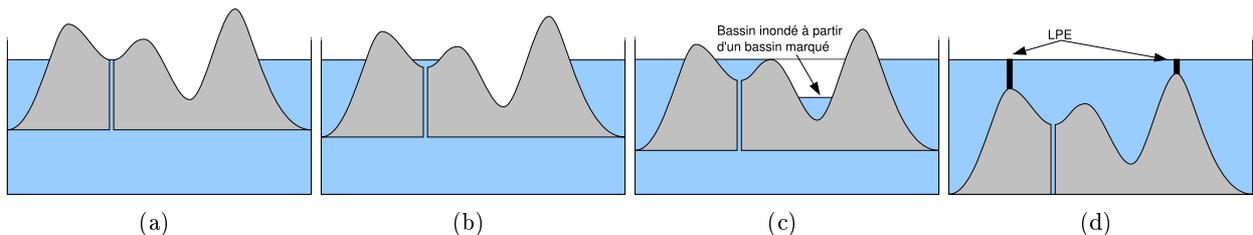


FIG. 5.5 – Processus d'immersion de la LPE avec des marqueurs.

Le problème sera donc de choisir correctement les marqueurs et nous proposons une méthode très simple mais très robuste pour choisir ces marqueurs dans le cadre de la segmentation des cocotiers.

5.2 Délimitation des cocotiers en imagerie Ikonos

5.2.1 Architecture de la méthode

Notre méthode comporte quatre étapes. La première étape consiste en un masquage des zones végétales. Parce que nous ne disposons pas de la bande proche infrarouge qui permet de calculer des indices de végétation pertinents nous avons dû utiliser une stratégie alternative basée sur une classification supervisée et une analyse de la texture. À l'issue de cette étape, les trois canaux RGB sont fusionnés en un seul canal grâce à une analyse en composantes principales (ACP), le traitement est poursuivi en utilisant cet unique canal. La seconde partie consiste en une construction des marqueurs tels que les maxima locaux et les minima locaux de l'image. Les maxima locaux sont les pixels candidats des centres des arbres. La troisième étape est le processus de segmentation, par ligne de partage des eaux, contraintes par les marqueurs précédemment calculés. La quatrième et dernière étape consiste à éliminer les régions aberrantes. Ces étapes sont présentées dans les sous-sections suivantes. La figure 5.6 résume les quatre étapes de la méthode.

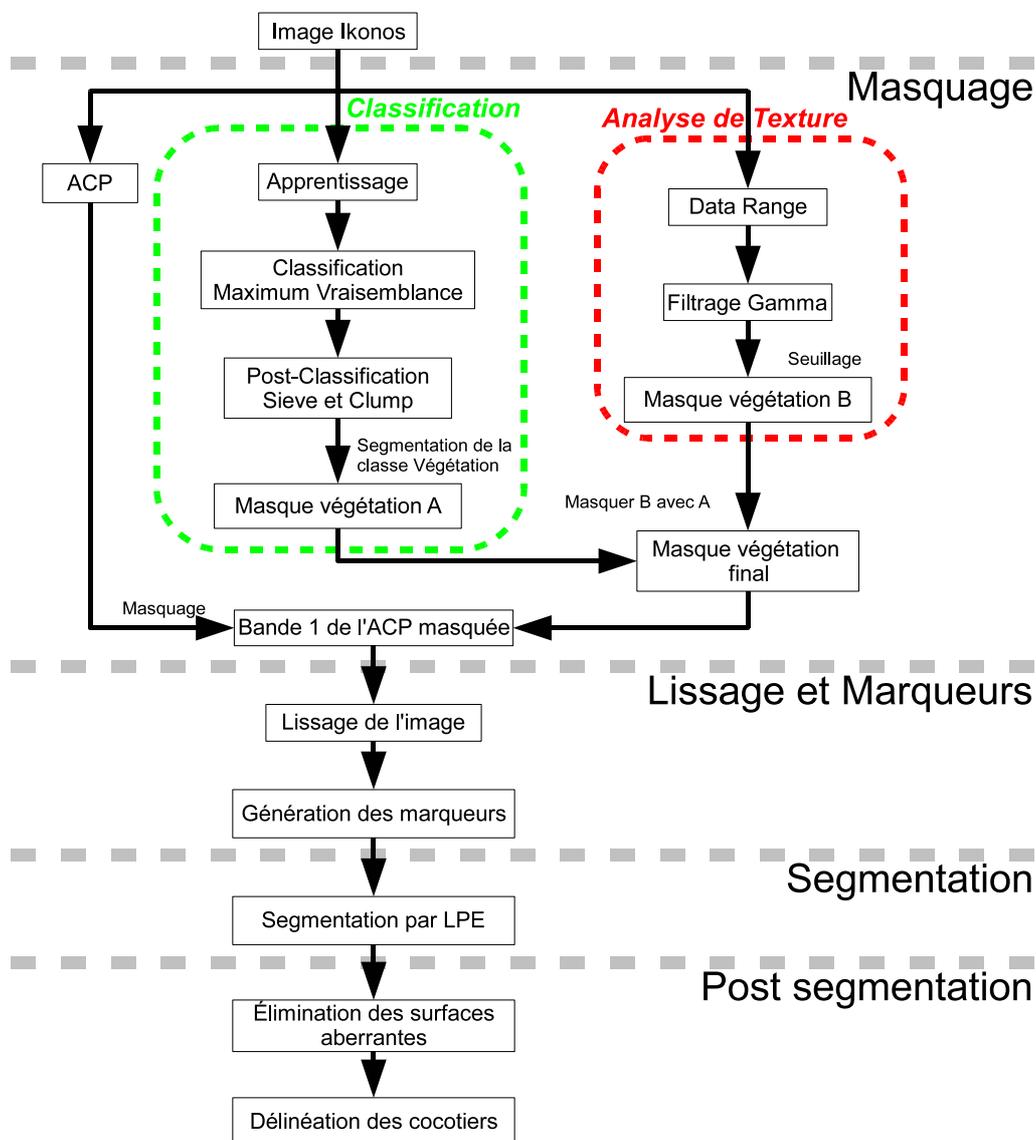


FIG. 5.6 – Diagramme de la méthode.

5.2.2 Masquage des zones végétales

Cette étape consiste à créer un masque des zones végétales, et plus précisément, des zones dans lesquelles nous souhaitons rechercher les arbres. La nature du sol n'est pas homogène : à certains endroits nous avons un sol entretenu, à d'autres endroits un sol non entretenu pouvant comporter du sable. La figure 5.7 montre les différentes natures du sol observées dans les images. Le sable réfléchit énormément de lumière, il est donc nécessaire de masquer au mieux les zones dont le sol est constitué de sable afin d'éviter des erreurs éventuelles de détection (c'est-à-dire détecter en tant que canopée une région n'étant que du sable).

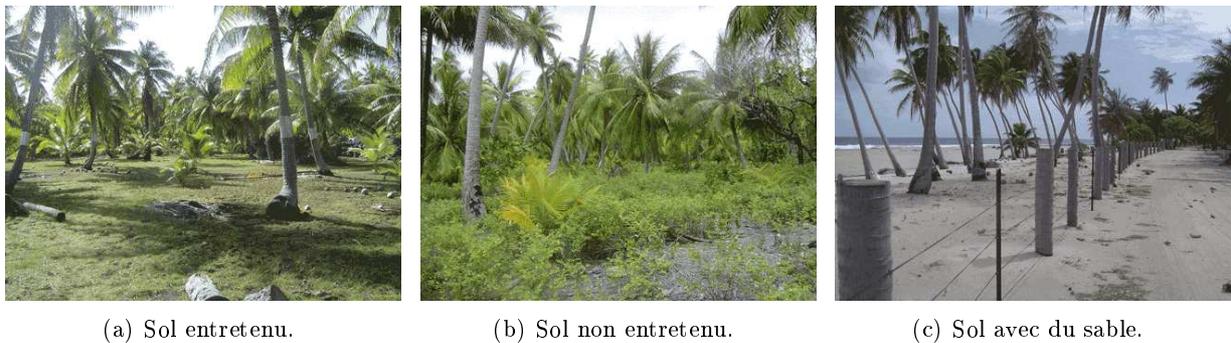


FIG. 5.7 – Exemples des différentes natures du sol.

La détection des zones végétalisées est effectuée par une classification bayésienne (détaillée dans la section 8.6). Pour cela, il est nécessaire de définir des régions d'apprentissage, puis la classification est réalisée avec les lois *a priori* estimées dans ces régions d'apprentissage. Le résultat de la classification bayésienne est régularisée à l'aide d'opérateurs morphologiques *sieve* et *clump*. L'opérateur *sieve* permet de retirer les pixels isolés d'une classe dont la population (exprimée en nombre de pixels) est inférieure à un seuil. L'opérateur *clump* est une fermeture morphologique de l'image de classification qui permet de combler une partie des pixels non classés. De cette manière, il devient aisé de segmenter la végétation du reste des objets présents dans l'image pour obtenir un premier masque de végétation. En plus des cocotiers présents dans les images, la présence de végétation basse est susceptible de perturber l'analyse. Il est donc nécessaire et important de réaliser une analyse de texture du couvert végétal pour séparer les différents motifs présents à l'aide du filtre de texture *Data Range*. La nouvelle image est ensuite filtrée avec le filtre adaptatif *Gamma* (avec une taille du noyau de convolution égal à une fenêtre de 7×7 pixels). Puis, il faut effectuer un seuillage de cette image lissée et la remasquer avec le premier masque de végétation obtenu précédemment. La valeur de seuil est obtenue en examinant un profil sur une portion de l'image contenant un mélange des différentes textures.

5.2.3 Fusion des bandes par ACP

Parallèlement aux opérations de masquage, une Analyse en Composante Principale (ACP) de l'image RGB originale est réalisée. L'ACP consiste à remplacer les trois composantes RGB, qui sont corrélés entre elles, par de nouvelles composantes : **Les Composantes Principales**. Ces nouvelles composantes sont des combinaisons linéaires des variables initiales ou attributs, à composantes non corrélées, de variance maximale et les plus liées en un sens aux attributs. La première composante contient l'ensemble de l'information [Ricotta and Avena, 1999]. Le tableau 5.1 montre les matrices de corrélations de l'image RGB et celle de l'ACP. Nous observons que dans l'image RGB, les canaux sont fortement corrélés alors qu'à la suite de l'ACP, les composantes principales ne sont plus corrélées. Le tableau 5.1(c) illustre bien le fait que la première composante de l'ACP contient l'ensemble de l'information avec une inertie de 94%.

La figure 5.8 montre l'image couleur originale, l'image par la transformée en ACP, le masque final des zones non-végétales et la première composante de l'ACP masquée de manière à ne garder que les zones de végétation haute.

	R	G	B
R	1		
G	0,93	1	
B	0,82	0,95	1

(a) Correlation (RGB)

	ACP1	ACP2	ACP3
ACP 1	1		
ACP 2	0,00	1	
ACP 3	0,00	0,00	1

(b) Correlation (ACP)

Bandes	Valeurs Propres
ACP 1	81149,45
ACP 2	4041,89
ACP 3	488,78

(c) Valeurs propres (ACP)

TAB. 5.1 – Comparaison des statistiques de l'image RGB et de l'ACP.

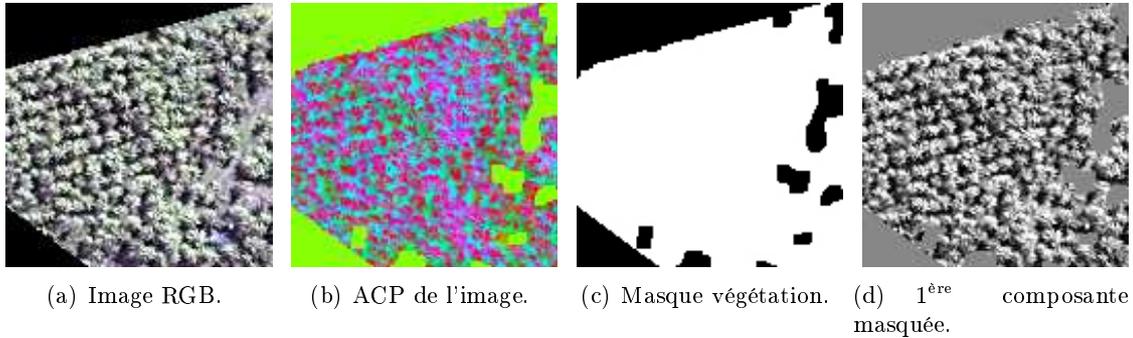


FIG. 5.8 – Masquage des zones d'intérêt.

En annexe C, le tableau C.1 récapitule toutes les valeurs des différents paramètres et les diagrammes illustrant chacune de ces étapes pour chaque *motu* traité.

5.2.4 Lissage des images

Après masquage des zones d'intérêt réalisé sur l'image fusionnée, l'image est lissée avec le filtre anisotrope de Perona & Malik (voir la section 4.1). Un tel filtre permet de lisser les zones homogènes tout en conservant les contours.

L'utilisation d'opérateurs (ouverture, fermeture, *TopHat* et *BottomHat*) de la morphologie mathématique permet d'accentuer les vallées (les pixels sombres) pour permettre d'améliorer la recherche des vallées entre les cocotiers proches dont les palmes se chevaucheraient. Les figures 5.9 présentent les effets du lissage sur les images Ikonos de cocotiers.

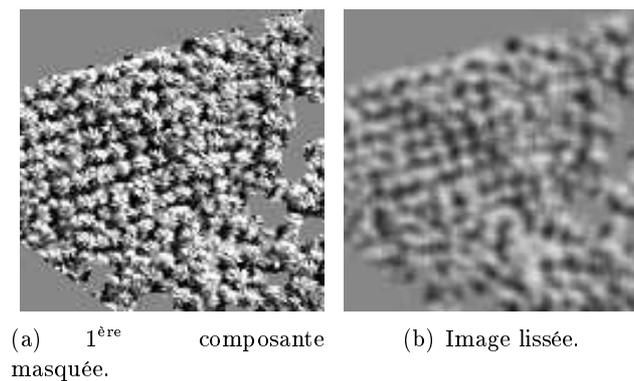


FIG. 5.9 – Lissage de l'image.

5.2.5 Génération des marqueurs

Pour obtenir une segmentation pertinente avec l'algorithme de ligne de partage des eaux, il est nécessaire d'initialiser correctement les points d'immersion grâce à des marqueurs. Sous les hypothèses décrites dans [Pollock, 1996], l'intensité lumineuse captée par le satellite est la plus forte près du sommet de l'arbre (voir la section 4.2). Les maxima locaux dans les zones de cocotiers correspondent

aux sommets des cocotiers. Les minima et les maxima locaux sont recherchés en utilisant la méthode décrite dans la section 4.4, c'est-à-dire par recherche des maxima et minima locaux dans les quatre directions. Ces derniers servent donc d'initialisation pour le processus d'immersion de la segmentation par ligne de partage des eaux.

La figure 5.10 montre le résultat de la recherche de ces marqueurs. Les pixels rouges correspondent au maxima locaux et les pixels jaunes sont le résultats du calcul du réseau des minima.

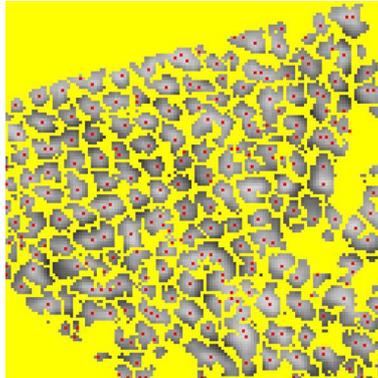


FIG. 5.10 – Marqueurs pour contraindre la segmentation par ligne de partage des eaux.

5.2.6 Post-segmentation

Les méthodes de ligne de partage des eaux ont l'inconvénient de fournir une sur-segmentation. Pour la limiter, on peut discriminer chaque région détectée selon leur surface afin d'obtenir une délimitation plus précise des cocotiers.

Sachant que la résolution au sol des pixels des données est de 1 m , la conversion d'une surface en pixel à une surface en m^2 se fait directement. Le houppier d'un cocotier se rapproche d'un cercle où peut être inscrit dans un cercle de rayon R (voir figure 5.11).

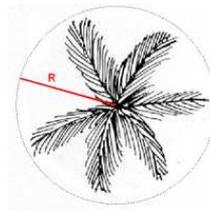


FIG. 5.11 – Canopée d'un cocotier inscrit dans un cercle de rayon R .

Pour estimer la surface du houppier d'un cocotier, il faut donc mesurer le rayon de ces cercles. Différentes surfaces sont calculées en fonction du rayon R . Le rayon R représente la longueur d'une palme projetée dans un plan parallèle au sol. À cause son propre poids, la palme d'un cocotier n'est pas droite mais courbée vers le sol, et donc la longueur projetée sur le plan est plus petite que la longueur réelle de la palme. Le tableau 5.2 montre différentes valeurs de surface pour différents rayon sachant que certaines valeurs sont improbables. Notamment seules les surfaces comprises entre 12 m^2 et 120 m^2 sont considérées. Les lignes grisées représentent les rayons d'un houppier qui ne seront pas gardés car la surface représentée est peu probable. Chaque région issue de la segmentation est mesurée en comptant les pixels qui se trouvent à l'intérieur et validée comme un houppier si la surface respecte les contraintes ci-dessus.

Et enfin, une fois le résultat de notre segmentation affiné, nous obtenons des régions qui représentent les couronnes des arbres contenus dans l'image. Dans la figure 5.12, nous montrons le résultat final obtenu après segmentation par ligne de partage des eaux et affinage. Les contours des arbres sont

rayon (en m)	surface (en m^2 ou en pixels)
1	3,14
2	12,56
3	28,26
4	50,24
5	78,5
6	113,04
7	153,86
8	200,96
9	254,34
10	314

TAB. 5.2 – Tableau donnant les surfaces (en m^2 ou pixel) en fonction du rayon (en m). Les lignes grisées indiquent un rayon improbable. Ces régions seront donc éliminées.

superposés en jaune sur l'image originale. Nous observons que la majorité des arbres sont correctement détectés. Les erreurs de détection sont principalement dues aux difficultés de générer un masque de végétation adéquat. Nous verrons dans le chapitre suivant comment quantifier cette erreur pour valider la méthode.

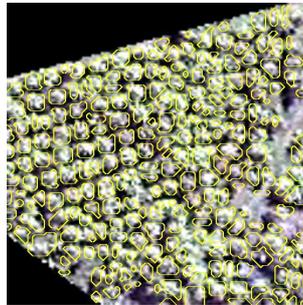


FIG. 5.12 – Post-segmentation : élimination des régions aberrantes.

5.2.7 Exploitation des résultats

À la fin du processus de délimitation des couronnes des arbres, chaque région représente la couronne d'un cocotier. À partir de cette information de surface et de positionnement de la couronne, nous pouvons extraire diverses informations telles que :

- la surface : pour toutes les régions, les surfaces associées à chacune d'entre elles sont mesurées. Ces surfaces sont exprimées en pixels mais nous avons vu dans la section 5.2.6 l'équivalence entre une surface exprimée en pixels et une surface exprimée en m^2 .
- le centre : pour chaque région, un barycentre est calculé. Ce barycentre est pondéré par les valeurs des pixels contenus à l'intérieur de la région en question. Ceci permet d'obtenir les coordonnées dans le référentiel de l'image de toutes les couronnes.
- Les couleurs moyennes : nous rappelons que nous avons en notre possession des images Ikonos à trois canaux (R,V et B). La couleur moyenne est calculée à l'intérieur de chacune région pour chaque canal. Cette information s'avère utile si on peut corrélérer la couleur d'une palme avec l'état de santé d'un cocotier par exemple.

Grâce aux coordonnées des centres des houppiers détectés, nous calculons les densités locales. Elles sont obtenues à partir du protocole suivant : pour chaque pixel de l'image, on dénombre le nombre d'arbres contenu dans le cercle de 12 m de rayon, centré sur ce pixel. Ce protocole a été établi dans [Jacq, 2006]. Il nous permet donc de comparer nos résultats avec les valeurs relevées sur le terrain par [Jacq, 2006] et par les missions terrain qui ont été réalisées dans le cadre de cette thèse. La figure 5.14 présente un exemple de carte de densité locale générée avec la méthode expliquée ci-dessus.

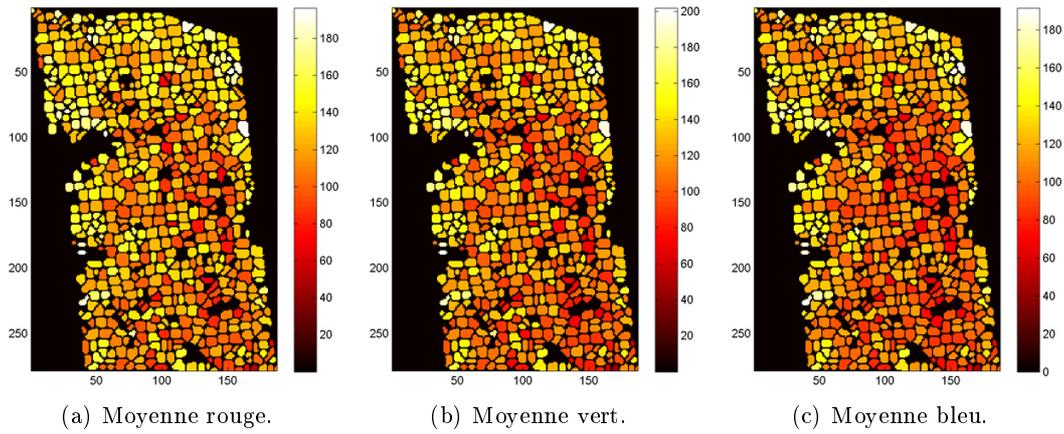


FIG. 5.13 – Exemples de couleurs moyennes des canopées extraites à partir des résultats de la segmentation des couronnes des arbres.

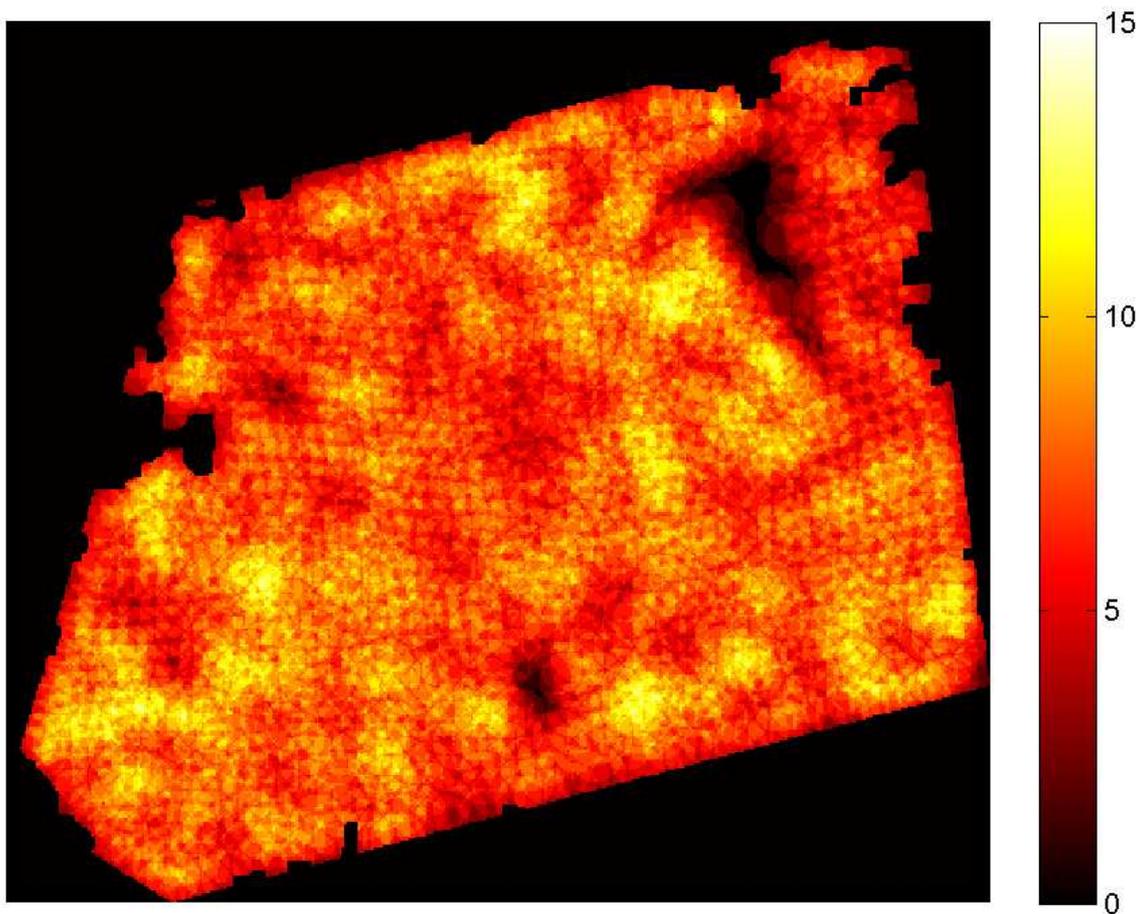


FIG. 5.14 – Exemple de carte de densité locale générée sur le secteur test 2 du motu principal.

Application à l'atoll de Tikehau et validation

La méthode de délimitation est appliquée à un ensemble de secteurs test situés sur le *motu* principal de Tikehau. Ces zones sont choisies de façon à représenter un grand nombre de configuration sur le terrain et à évaluer ainsi la robustesse de la méthode. Un protocole de validation en deux temps, validation par interprétation visuelle et validation sur le terrain, est proposé. Une fois la méthode validée, elle est appliquée à l'ensemble des *motu* de l'atoll.

6.1 Résultats et Validation

6.1.1 Résultats sur des zones de test

De façon générale, sur les données satellites que nous avons en notre possession, nous distinguons trois types de cocoteraies :

- les plantations artificielles,
- les plantations naturelles,
- les plantations mixtes.

Les plantations artificielles sont très distinguables par rapport aux autres types de plantations car elles sont caractérisées par un alignement et un espacement régulier entre chaque cocotier.

Les plantations naturelles, au contraire, ne montrent pas d'organisation spatiale spécifique, c'est-à-dire que dans ce type de plantation nous n'observons aucun alignement ou espacement régulier entre chaque cocotier. Chacun de ces cocotiers pousse en compétition avec ses voisins, et donc dans ce type de plantation nous pouvons observer des chevauchements de palmes entre cocotiers adjacents.

Les plantations mixtes montrent qu'il y a eu une certaine organisation dans la cocoteraie, mais qu'ensuite celle-ci a dû être laissée à l'abandon. Au fil du temps, des cocotiers ont poussé un peu partout dans les espaces de l'ancienne plantation.

Pour étudier la pertinence et la robustesse de notre méthode de comptage des cocotiers, nous l'appliquons sur différents secteurs test de taille variée (de 41000 m² à 129000 m²) du *Motu #16* (voir la figure 3.4) de Tikehau. Ce *motu* a l'avantage de comporter une grande variété de cocoteraies (naturelles, artificielles ou mixtes).

Le tableau 6.1 présente les résultats obtenus sur les secteurs de test du motu 16. Pour chacun des secteurs test, il est donné sa surface, le nombre d'arbres détectés, la surface moyenne de chacune des couronnes des arbres et pour chaque canal (rouge, vert et bleu) la couleur moyenne des couronnes. La pertinence et la robustesse seront ensuite évaluées grâce à deux procédés de validation.

6.1.2 Validation par photo interprétation humaine

Une première vérité image est établie par photo interprétation humaine (le protocole est présenté en annexe E). Un opérateur interprète les images et marque si chaque structure est un cocotier ou non. Ceci permet d'obtenir une estimation de l'erreur de détection de notre méthode. Ces vérités photos sont établies sur les secteurs test du motu principal de Tikehau et fournissent une première validation

Secteurs	Surf.(m ²)	# arbres	Surf.Moy.	Image couleur (pixel codé sur 8 bits par canal)		
				Canal Rouge	Canal Vert	Canal Bleu
1	128824	908	55,09	Moyenne	Moyenne	Moyenne
2	112897	986	51,68	122,59	120,36	114,44
3	63029	495	51,57	130,94	130,54	128,46
4	41194	346	53,91	130,23	126,39	120,38
5	58886	436	52,99	125,52	122,16	112,32
6	128044	1269	50,89	146,20	146,74	147,74
				146,11	147,75	153,36

TAB. 6.1 – Résultats obtenus sur chacun des secteurs tests.

de la méthode utilisée pour la détection des houppiers et de leur centre. Le tableau 6.2 présente pour chaque secteur de test le nombre de cocotiers détectés, le nombre d'arbres interprétés par l'être humain et ainsi que le pourcentage d'erreurs entre ces deux résultats.

Secteur	Nb Method	Interp.	% erreur
1	908	818	9,91%
2	986	912	7,51%
3	495	450	9,09%
4	346	356	-2,89%
5	436	399	8,49%
6	1269	1198	5,59%
Total	4440	4133	6.81%

TAB. 6.2 – Tableau du résultat de la vérité photo par interprétation humaine.

Les résultats de la photo interprétation humaine montre que la méthode détecte les cocotiers visibles (discernables sur l'image) avec une erreur inférieure à 10%. Ces résultats sont néanmoins à relativiser car la méthode d'interprétation humaine a des limitations : difficultés à reconnaître un cocotier, la fatigue de l'oeil qui s'habitue aux motifs présents dans l'image et induit une difficulté d'interprétation des objets. La méthode possède aussi un biais dû aux effets de bords (tous les arbres en contact avec le bord de l'image ne sont pas pris en compte) qui diminue quand la taille des secteurs traités augmente, à la perte d'arbres isolés, aux agglomérats, ...

6.1.3 Validation par vérité terrain

Protocole utilisé

Une mission terrain sur le motu principal de Tikehau a été organisée dans le cadre de cette thèse pour établir une vérité terrain. Comme dans la sous-section 5.2.7, nous nous sommes basé sur le protocole défini par [Jacq, 2006]. Le protocole diffère légèrement de celui utilisé pour générer les cartes de densités à partir des images. Nous avons positionné des placettes de 12 m de rayon agencées en quinconce sur les secteurs de test. La taille L de la maille varie afin de respecter un pourcentage minimal de l'échantillonnage. Pour chacune de ces placettes, nous comptons les cocotiers afin d'obtenir une densité locale. La figure 6.1 illustre la disposition en quinconce des placettes.

Ceci permet d'obtenir pour chaque secteur un taux d'échantillonnage supérieur ou égal à 2,8%. Le tableau 6.3 récapitule pour chaque secteur :

- la surface du secteur,
- la taille L de la maille,
- le nombre de placette,
- la surface échantillonnée,
- et le pourcentage d'échantillonnage.

Comme nous pouvons le voir, la surface totale des secteurs est de 532870 m², sur lesquels nous avons effectués des mesures pour 76 placettes. Ces placettes représentent une surface échantillonnée de 34352 m², soit 6.45% de la surface totale des secteurs de test.

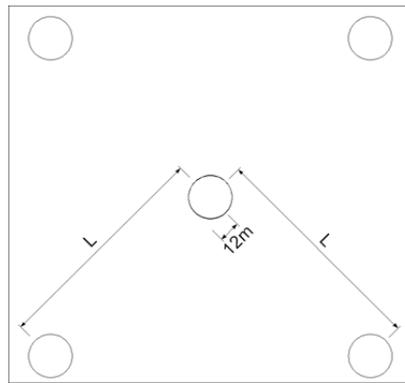


FIG. 6.1 – Dispositif des placettes en quinconce : chaque placette est un cercle de 12 m de rayon.

Secteur	Surface en m ²	Taille maille en m	Nb Placette	Surf. échan. en m ²	% échan.
1	128820	125	8	3616	2,81%
2	112900	125	10	4520	4,00%
3	63030	100	6	2712	4,30%
4	41190	50	24	10848	26,34%
5	58890	100	7	3164	5,37%
6	128040	100	21	9492	7,41%
Total :	532870		76	34352	6,45%

diamètre placette : 12 m
surface : 452 m²

TAB. 6.3 – Tableau récapitulatif du nombre de placettes par secteur.

Difficultés rencontrées

Lors de la mission terrain, deux principales difficultés ont été rencontrées : le corail soulevé et les strates des cocotiers dans les plantations.

Au sud du motu principal, côté océan près de l'aéroport, il y a du corail soulevé¹, appelé *FEO*. Sur les images nous ne pouvons pas distinguer ce type de terrain. Mais, une fois sur place, nous avons été confronté à des difficultés d'accès à ces endroits (voir figure 6.2). Il a donc été impossible d'effectuer le comptage à l'intérieur de placettes contenant du corail soulevé. D'autres placettes, situées en dehors de zone de corail soulevé, ont été choisies afin de respecter l'échantillonnage minimal prévu.



FIG. 6.2 – Difficulté d'accès dans les zones de corail soulevé.

Dès la première visite terrain effectuée dans les cocoteraies, il s'est avéré plus judicieux de différencier les cocotiers en fonction de leur emplacement par rapport aux cocotiers voisins. En effet, il est difficile de distinguer dans une image satellite les cocotiers ne se trouvant pas en canopée. C'est-à-dire que sur le terrain il y a des cocotiers plus petits dont la canopée est partiellement masquée par un cocotier voisin plus grand. Nous avons donc classifié les cocotiers en trois strates :

- Strate 1 : les cocotiers en canopée, non masqués par des cocotiers voisins,
- Strate 2 : les cocotiers dont les palmes se chevauchent,
- Strate 3 : les cocotiers partiellement ou totalement masqués par un cocotier plus grand.

¹Soulèvement de la plaque corallienne sous l'effet du poids du cratère lors de l'effondrement de celui-ci.

La figure 6.3 présente un schéma des différentes strates (avec une vue de profil ainsi qu'une depuis la verticale). La figure 6.4 illustre les différentes strates avec des exemples observés sur le terrain.

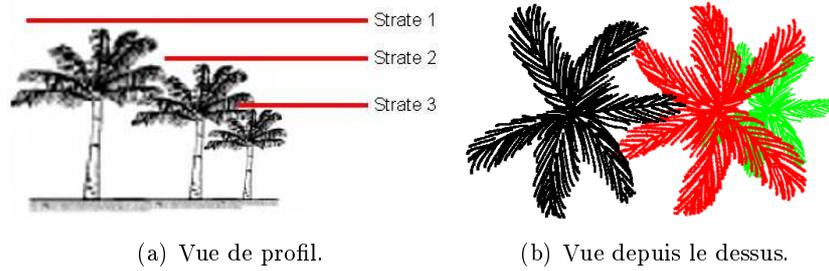


FIG. 6.3 – Schéma des différentes strates présentes sur le terrain. À gauche, la figure représente une vue de profile, et à droite une estimation de la vue depuis le dessus.



FIG. 6.4 – Exemples de différentes strates observées sur le terrain.

La classification sur le terrain d'un arbre dans l'une des trois strates choisies est soumise à une interprétation humaine ce qui induit un biais. De ce fait, lors des mesures sur le terrain, cette classification doit se faire uniquement par la même personne afin de respecter le même biais.

6.1.4 Validation de la méthode sur les zones test du motu 16

La méthode est validée en confrontant les résultats obtenus avec la vérité photo (sous-section 6.1.2). Les données recueillies sur le terrain sont ensuite utilisées pour établir la corrélation entre le nombre de cocotiers recensés sur le terrain avec la carte de densité générée à la suite de détection des houppiers et de leur centre. Pour placer les cartes de densité et les recensements terrains des cocotiers dans le même référentiel, nous utilisons le GPS en tenant compte de la précision donnée par le GPS lors de la prise de points de référence sur le terrain. Ainsi, sur notre image de densité, nous regardons dans un cercle de rayon R_{GPS} correspondant à la précision du GPS, pour un point de référence donné, les valeurs minimales et maximales. De cette manière, nous obtenons une correspondance entre ce que nous avons sur le terrain et les résultats obtenus avec notre méthode.

Nous nous sommes intéressés aux cocotiers classifiés comme appartenant à la strate $S1$. Parmi les lignes de des différents tableaux données en annexe F, nous remarquons que le nombre de cocotiers détectés présents sur la carte de densité est parfois supérieur. Ceci est principalement dû au fait que des arbres classifiés en tant que type de strate $S2$ sont détectés, ceci peut s'expliquer par le fait que la canopée d'un arbre de type $S2$ n'est que très partiellement couvert par son voisin de type $S1$, et donc ces arbres sont différenciables. Considérons les trois hypothèses suivantes :

- H1 : seuls les cocotiers de la classe $S1$ sont détectables sur l'image,

- H2 : seuls les cocotiers des classes S1 et S2 sont détectables sur l'image,
- H3 : les cocotiers des trois classes S1,S2 et S3 sont détectables.

L'hypothèse H1 est la plus probable car les arbres situés en canopée bénéficient d'un ensoleillement optimal, et donc rentrent dans le cadre de l'illumination décrite sur la figure 4.4. Les hypothèses H2 et H3 sont fortement improbables, même si quelques arbres de la strate S2 peuvent bénéficier d'un ensoleillement suffisant pour les rendre détectables. Nous établissons, pour ces trois hypothèses, les statistiques suivantes :

- le taux de cocotiers détectés (DE) comme le rapport du nombre de cocotiers détectés sur le nombre total de cocotiers recensés sur le terrain,
- le taux de cocotiers non détectables (NDA) comme le rapport entre les cocotiers non détectables (en fonction de l'hypothèse) et le nombre total de cocotiers recensés sur le terrain,
- le taux de cocotiers non détectés (NDE) comme le pourcentage entre les cocotiers non détectés (parmi les cocotiers détectables) et le nombre de cocotiers total trouvé sur le terrain,
- le taux de détection pondéré (TDP) comme le pourcentage de cocotiers détectés sur le nombre de cocotiers détectables, ceci permet de s'affranchir des cocotiers non détectables et d'avoir un chiffre plus représentatif de la précision de la méthode.

Nous remarquons que $DE + NDA + NDE = 100\%$, le nombre de cocotiers non détectables étant fixe, nous pouvons, dans un esprit d'amélioration de la méthode, essayer d'augmenter le taux DE et donc de baisser le taux NDE . Un taux NDE de 0% signifie que l'on a détecté tous les cocotiers présents sur le terrain, +25% signifie que l'on a détecté 25% de cocotiers en plus, et -30% signifie que l'on en a détecté 30% en moins que ce qu'il y a réellement sur le terrain.

Le tableau 6.4 récapitule les résultats de la mise en correspondance entre les données extraites à partir de la méthode de segmentation des houppiers et celles du comptage systématique faites sur le terrain. On trouvera en annexe F les résultats concernant tous les secteurs de test (voir tableaux F.1,F.2, F.3, F.4 et F.5).

	H3 DE (S1+S2+S3)			NDA (S2+S3)	H1 NDE			TDP (S1)			NDA (S3)	H2 NDE			TDP (S1+ S2)		
	max	min	moy.		max	min	moy.	max	min	moy.		max	min	moy.	max	min	moy.
	Secteur 02 :	84%	40%		61%	41%	19%	-25%	-2%	142%		67%	103%	14%	46%	46%	25%
Secteur 03 :	93%	43%	69%	32%	25%	-25%	-1%	137%	63%	102%	4%	54%	54%	27%	96%	44%	72%
Secteur 04 :	102%	42%	71%	27%	31%	-29%	1%	140%	58%	98%	3%	56%	56%	26%	104%	43%	73%
Secteur 05 :	68%	27%	46%	35%	37%	-3%	19%	105%	42%	71%	19%	54%	54%	36%	83%	34%	56%
Secteur 06 :	83%	37%	60%	43%	20%	-26%	-3%	146%	66%	105%	9%	53%	53%	31%	92%	41%	66%
Motu16	86%	38%	61%	37%	26%	-22%	2%	135%	59%	96%	10%	53%	53%	29%	95%	42%	68%
sans 05	90%	40%	64%	37%	23%	-27%	-1%	163%	73%	124%	8%	52%	52%	28%	102%	46%	76%

TAB. 6.4 – Tableau récapitulatif de la mise en correspondance des données terrain avec les cartes de densités générées à partir des résultats de détection de la méthode sur l'ensemble des secteurs de test du Motu 16.

Dans le tableau 6.4, l'hypothèse H3, qui considère que nous pouvons détecter tous les cocotiers dans l'image, est totalement improbable mais permet d'estimer le taux de détection global DE de la méthode. Les résultats donnent une estimation de 61% en moyenne sur tout le motu principal, c'est-à-dire un deux tiers des cocotiers : ce chiffre est faible car il ne tient pas compte du fait que certains cocotiers ne sont pas discernables dans les images (strate S3), et que d'autres cocotiers sont agglomérés car trop proches de leurs voisins (strate S2). Néanmoins ce chiffre peut servir à estimer le nombre de cocotiers au total en fonction du dénombrement obtenu avec la méthode. Par exemple, si on estime 986 cocotiers dans le secteur 2, on peut donc en déduire qu'il y aurait 1616 cocotiers en tout sur ce secteur. Pour les secteurs dont nous ne disposons pas de la valeur du TD, nous utilisons le TD moyen (61%), on trouve ainsi 1490 cocotiers pour le secteur 1 (voir le tableau 6.5).

La valeur du TDP du cas H1 est important, il représente l'efficacité de la méthode. En effet, il est impossible d'avoir un DE (H3) égal à 100% puisqu'un certain nombre d'arbres ne sont pas visibles. Notre méthode ne pouvant évidemment pas détecter des arbres invisibles dans l'image, le chiffre qu'il faut donc optimiser est le TDP de H1. Si l'on regarde les chiffres, la méthode détecte globalement 96% des cocotiers théoriquement discernables sur l'image. Idéalement elle devrait détecter 100% mais ce serait compter sans les limitations de la méthode : problème des masques, problème des effets de

bord, effet d'agglomérations qui rendent la surface de certaines couronnes trop importantes et ainsi sont rejetées en post-traitement, la présence de nuages, etc.

Secteur	Nb méthode	TD	Nb estimé
1	908	61%	1490
2	986	61%	1616
3	495	69%	714
4	346	71%	485
5	436	46%	952
6	1269	60%	2112

TAB. 6.5 – Extrapolation des résultats sur les zones de test du motu 16.

Il faut aussi optimiser le TDP de H2 car c'est la configuration la plus présente sur le terrain. Si l'on regarde les chiffres, la méthode détecte entre 96% et 68% des cocotiers directement discernables sur l'image.

6.1.5 Comptage exhaustif sur un *motu*

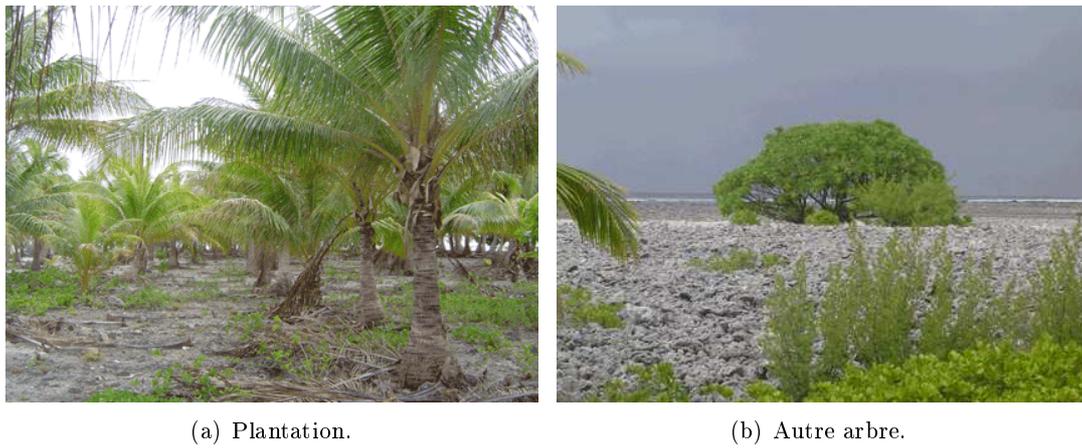
Pour vérifier la validité de la stratégie d'échantillonnage, nous effectuons un comptage exhaustif de tous les cocotiers sur un espace restreint. Le *motu* 25 (voir figure 6.5)), dont la surface est suffisamment petite pour rendre l'opération possible, présente la particularité d'avoir des cocotiers plantés (depuis 1986) avec une densité plus importante (la distance moyenne entre deux arbres est de 4 m) que sur les zones du *motu* principal, dont les arbres ont été principalement plantés après le cyclone de 1982 par le SDR avec une distance moyenne de 8 m. Ce cas de figure se rapproche de celui des *motu* isolés dont les cocotiers poussent de manière anarchique (pas de gestion de l'espacement entre deux arbres). Lors du comptage exhaustif des cocotiers présents sur ce *motu*, nous avons relevé les différentes strates auxquelles appartiennent les arbres. Ce type de cocoteraie s'avère être plus complexe que ceux trouvés sur le *motu* principal (*motu* 16). La figure 6.6(a) montre le type de plantation rencontrée sur le *motu* 25 qui comporte également des arbres qui ne sont pas des cocotiers (figure 6.6(b)).



FIG. 6.5 – Le Motu 25.

Après pré-traitements et masquage des zones, la méthode a détecté 192 cocotiers. La figure 6.7 montre les résultats obtenus pour la détection des cocotiers. La figure 6.8 montre les moyennes des canaux RGB pour chaque partie segmentée. Les effets de bords des résultats sont importants car la zone traitée est petite.

Le tableau 6.6 indique que la méthode détecte 108% des arbres visibles (8% de plus principalement dû aux problèmes de générer correctement le masque de végétation) sur l'image (contre 96% sur le *motu* principal); et 68% du nombre d'arbres effectivement présents sur le *motu* (contre 61% sur le *motu*



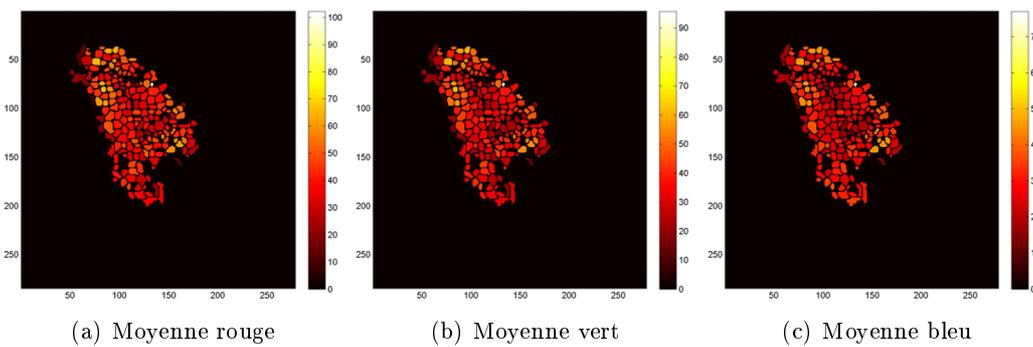
(a) Plantation.

(b) Autre arbre.

FIG. 6.6 – Photos du Motu 25.



FIG. 6.7 – Résultats en image du motu 25.



(a) Moyenne rouge

(b) Moyenne vert

(c) Moyenne bleu

FIG. 6.8 – Moyenne des canopées détectées dans chacun des trois canaux Rouge, Vert et Bleu.

Strates	S1	S2	S3	Nb détectés	H3		H1		H2		
					DE (S1+S2+S3)	NDA (S2+S3)	NDE	TDP (S1)	NDA (S3)	NDE	TDP (S1+S2)
Nb arbres	177	55	50	192	68%	37%	-5%	108%	18%	14%	83%
Total	282										

DE Detectés
 NDA Non Detectables
 NDE Non Detectés
 TDP Taux de Detection Pondéré

TAB. 6.6 – Résultat du comptage exhaustif des cocotiers présents sur le motu 25.

principal). Le taux d'arbres non détectables est de 37% (même chiffre que sur le motu principal) et le taux d'arbres non détectés est de -5% (contre 39% sur le motu principal), ce qui implique une perte

de 9%. En considérant l'hypothèse H2, nous constatons que le taux de détection pondéré est de 83% contre 76% pour le *motu* principal. La similarité des chiffres entre notre méthode exhaustive et notre méthode par échantillonnage systématique nous permet de valider notre procédure terrain.

6.2 Traitement de l'ensemble des *motu*

Après avoir validé la méthode de dénombrement des cocotiers sur le *motu* principal, nous l'appliquons à l'ensemble des *motu* constituant l'atoll de Tikehau et ceci pour montrer que notre approche peut traiter beaucoup de données en des temps raisonnables et être utilisée en production. Ces résultats ont été communiqué au Service du Développement Rural pour l'exploitation des résultats. Nous présentons deux cas (celui du *motu* 13 et celui du *motu* 16) pour lesquels les images sont grandes et qui ont nécessité un découpage en imageries des zones à traiter. Les résultats du dénombrement des cocotiers sur l'ensemble des *motu* sont ensuite présentés.

6.2.1 Traitement du *motu* au Nord-Est (*motu* #13)

Comme le montre la figure 6.9, une très grande partie de l'image est constituée d'eau. Afin de réduire les temps de calcul sur cette image (13654×10606 pixels soit un espace mémoire utilisé sur le disque de 828 Mo), nous avons divisé le *motu* en plusieurs sous images et éliminé ainsi les zones d'eau. Nous avons choisi de traiter ce *motu* de cette manière parce que nous arrivons à des limites techniques en termes d'utilisation de l'espace mémoire utilisé par la méthode.

Des essais préalables ont été effectués sur la zone 1 du *motu* 13. Nous avons constaté qu'en traitant entièrement cette zone, la méthode détectait 8492 cocotiers et qu'en traitant la même zone avec les imageries, la méthode obtenait un total de 8477 cocotiers ; ce qui fait une perte marginale de 15 arbres. On peut créer des imageries qui se chevauchent légèrement afin de ne pas perdre trop d'arbres lors de la détection mais aussi parce que nous rappelons que lors de la segmentation nous retirons les zones détectées en contact avec le masque ou le bord de l'image afin d'éviter d'éventuels effets de bord.

La figure 6.9 illustre la méthode de subdivision en imageries du grand *motu* au Nord-Est de l'atoll de Tikehau. Ce découpage en imageries des *motu* reste largement supervisé dans notre étude mais il est parfaitement possible de l'automatiser.

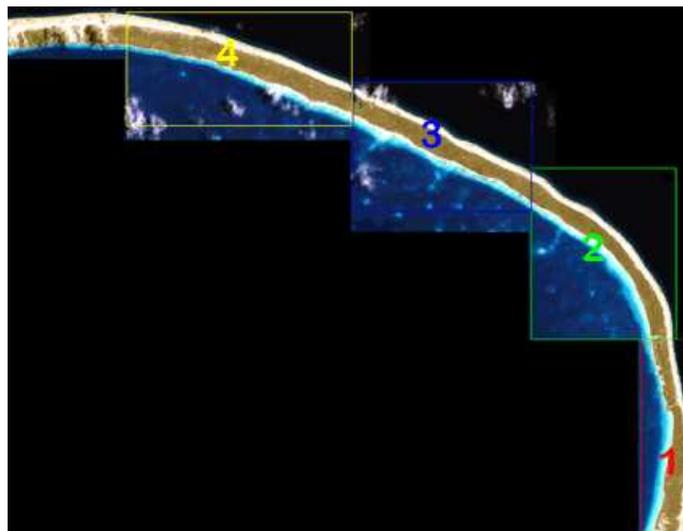


FIG. 6.9 – Présentation du grand *motu* au Nord-Est de l'atoll de Tikehau.

Même en effectuant une telle subdivision de l'image, il rest encore beaucoup de zones maritimes et nous avons encore dû subdiviser chaque partie. La figure 6.10 illustre comment cette subdivision est réalisée. Les autres subdivisions du *motu* 13 sont visibles à la figure G.1 de l'annexe G.

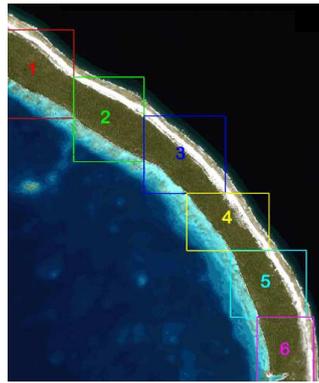


FIG. 6.10 – Subdivision des images.

Dans le tableau 6.7, sont consignés les résultats de la détection des cocotiers sur l'ensemble du *motu*. Nous pouvons voir que la méthode a détecté 48623 cocotiers sur une surface végétale de 537,852 ha, soit une densité de 90 cocotiers par hectare. Un tableau similaire pour chacune des sous-parties du *motu* est disponible à l'annexe G dans le tableau G.1.

Statistiques pour le Motu 13					
Nb cocotiers	48623				
Surface végétale	5378520 m ²	537,852 ha			
Statistiques des surfaces					
moy.	49,55	std	22,14	var	491,23
Statistiques des couleur					
Bande R	moy.	58,21	std	7,66	var 62,57
Bande V	moy.	35,26	std	6,10	var 60,44
Bande B	moy.	29,77	std	5,32	var 28,83

TAB. 6.7 – Tableau des résultats sur l'ensemble du *motu* 13.

6.2.2 Traitement du *motu* principal (*motu* #16)

Le problème de subdivision du *motu* se présente aussi ici. Nous avons donc utilisé la même méthode de division pour le traitement de ce *motu*. La figure 6.11 présente la subdivision utilisée. La visualisation de chacun de ces parties avec les résultats obtenus se trouvent dans la section G.2 de l'annexe G.

Après mise en réunion des résultats obtenus pour chaque partie (voir le tableau G.2 pour les détails), le tableau 6.8 présente les résultats de la détection des cocotiers. Nous remarquons que sur la surface végétale traitée de 220.24 ha, la méthode a détecté 13571 cocotiers, ce qui donne une densité globale de 60 cocotiers par hectare.

Statistiques pour le Motu 16					
Nb cocotiers	13571				
Surface végétale	2202384 m ²	220,24 ha			
Statistiques des surfaces					
moy.	55,62	std	24,52	var	601,87
Statistiques des couleur					
Bande R	moy.	55,34	std	10,61	var 113,27
Bande V	moy.	50,82	std	10,80	var 117,02
Bande B	moy.	40,77	std	9,06	var 82,84

TAB. 6.8 – Tableau des résultats sur l'ensemble du *motu* 16.



FIG. 6.11 – Présentation du *motu* principal de Tikehau avec les quatre parties traitées.

6.2.3 Résultats sur l'ensemble des *motu*

Nous avons traité les dix huit autre *motu*, qui ne comportent pas les difficultés rencontrées sur les *motu* 13 et 16. Le tableau 6.9 résume les résultats obtenus en indiquant la superficie des zones végétales dans lesquelles la méthode réalise la détection ainsi que le nombre d'arbres détectés. Rappelons que les superficies données ne sont en aucun cas celles des *motu* mais celles des masques uniquement. Pour chaque *motu* traité, la densité moyenne est aussi donnée en *arbres/ha*. Nous constatons finalement que pour une superficie totale de 1052,7 *ha*, la méthode a détecté 82444 cocotiers. Ce chiffre ne représente pas le nombre exact de cocotiers plantés sur l'ensemble des *motu* de l'atoll de Tikehau en raison des difficultés rencontrés lors de la génération des différents masques, mais compte tenu de la mise en correspondance des relevés sur le terrain présenté dans la section 6.1, ce chiffre est une bonne estimation de la cocoteraie de l'atoll Tikehau dans sa globalité. Le tableau 6.9 présente les résultats obtenus sur l'ensemble des *motu* de Tikehau avec une estimation de la réalité en prenant un TD moyen de 61%.

6.2.4 Cas des zones d'ombres induites par des nuages de basse altitude

Un des problèmes récurrent constaté sur les image est la présence de nuage ainsi que leur ombre portée sur les *motu* sur nos images. Comme nous pouvons le voir sur la figure 3.2, certains *motu* situés au Sud de l'atoll ont une couverture nuageuse très importante qui nous empêche de traiter ces zones. Une couverture nuageuse partielle est également présente sur certains *motu* et pour lesquels nous obtenons une erreur de classification des zones végétales lors de la création du masque de la végétation haute après l'analyse de texture. Ceci produit une erreur de l'estimation de la surface totale de la végétation haute. La figure 6.12 illustre un exemple d'un nuage présent dans l'image avec son ombre portée sur une zone végétale. La classe "Végétation" issue du résultat de la classification bayésienne est superposée en vert ainsi que le contour final de la végétation haute dans laquelle la dénombrement des cocotiers sera effectuée et on constate clairement que les zones d'ombres sont exclues de la segmentation. Néanmoins, notre méthode fonctionne correctement sur les zones d'ombres comme le prouve la figure 6.13 qui montre les résultats de la segmentation et de la détection des cocotiers dans cette région.

Pour remédier à ce problème, il suffit donc d'inclure les zones d'ombres dans l'étape de masquage. On peut définir une nouvelle classe ayant les caractéristiques spectrales d'une surface végétale avec une ombre et de reclassifier l'ensemble mais malheureusement, le classifieur bayésien ne parvient pas à attribuer correctement la bonne classe aux pixels des zones d'ombres. La solution est de traiter

Motu	Surface végétale		# arbres	Densité	TD (moyen)	Nb estimés
	en m ²	en ha				
1	229 970	22,997	2 100	91,32	61%	3 443
2	84 268	8,4268	307	36,43	61%	503
3	133 498	13,3498	1 224	91,69	61%	2 007
4	166 585	16,6585	568	34,10	61%	931
5	272 485	27,2485	1 185	43,49	61%	1 943
6	184 600	18,46	675	36,57	61%	1 107
7	NON TRAITÉ, SURFACE < 1 Ha					
8	NON TRAITÉ, SURFACE < 1 Ha					
9	NON TRAITÉ, SURFACE < 1 Ha					
10	294 064	29,4064	455	15,47	61%	746
11	398 705	39,8705	2 123	53,25	61%	3 480
12	484 723	48,4723	2 267	46,77	61%	3 716
13	5 378 520	537,852	48 623	90,40	61%	79 710
14	299 779	29,9779	1 980	66,05	61%	3 246
15	72 338	7,2338	713	98,57	61%	1 169
16	2 202 384	220,2384	17 571	79,78	61%	28 805
17	25305	2,5305	234	92,47	61%	384
18	13151	1,3151	113	85,93	61%	185
19	40502	4,0502	437	107,90	61%	716
20	43153	4,3153	405	93,85	61%	664
21	NON TRAITÉ, SURFACE < 1 Ha					
22	97204	9,7204	631	64,92	61%	1 034
23	NON TRAITÉ, SURFACE < 1 Ha					
24	94924	9,4924	740	77,96	61%	1 213
25	10753	1,0753	93	86,49	61%	152
Total :	10 526 911	1052,6911	82 444	78,32		
				Total estimé : 135 154		

TAB. 6.9 – Tableau récapitulatif du résultat de la détection des cocotiers sur l'ensemble des motu de Tikehau.

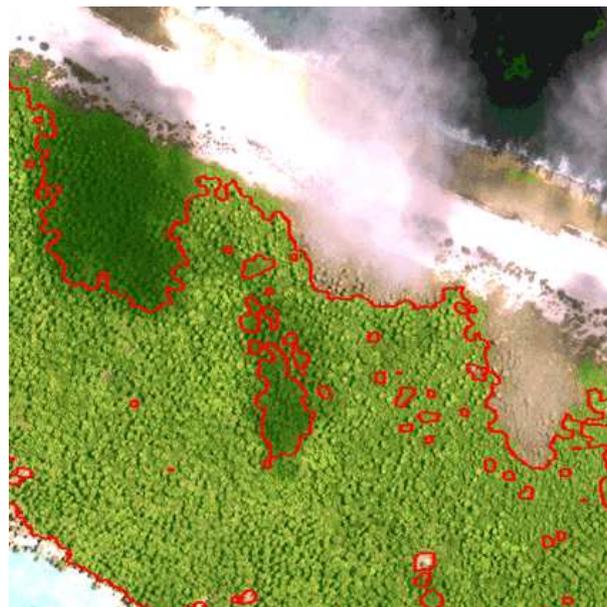


FIG. 6.12 – Exemple de couverture nuageuse : en vert la classe "Végétation" et en rouge le contour du masque de la végétation haute généré après l'analyse de texture.

séparément les zones d'ombres en les segmentant. Ceci est possible avec des techniques de seuillage dans des espaces de représentation des couleurs appropriés (voir la sous-section 7.1.1). La connaissance des zones d'ombres permet également de quantifier indirectement la hauteur des arbres et peut être utilisé pour affiner le type de population de cocotiers observés dans les images. Pour cette raison, nous traitons le cas des zones d'ombres et leur exploitation dans le chapitre suivant, à la marge de la problématique traitée dans cette partie. Finalement, nous constatons que la méthode permet, en modifiant seulement le calcul des masques, d'obtenir des résultats très satisfaisants dans les zones d'ombres.

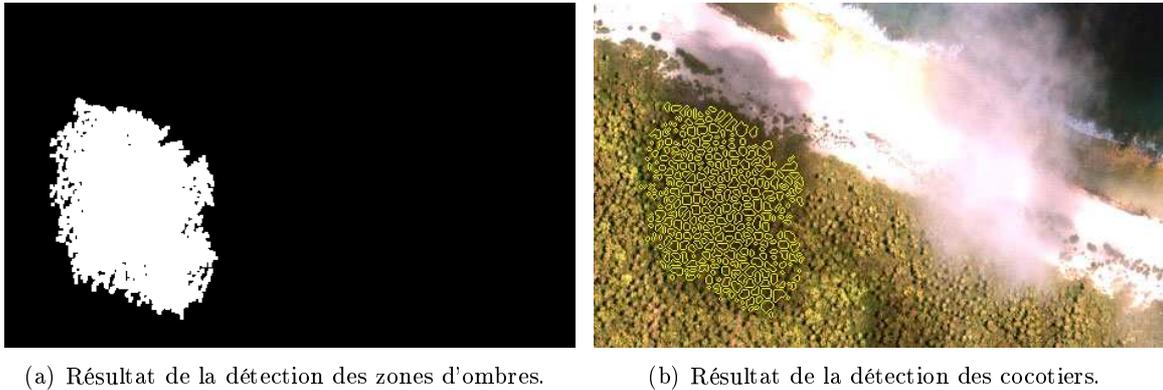


FIG. 6.13 – Exemple de traitement séparé pour les zones d'ombres.

Une autre erreur est produite lors de la génération des masques de la végétation haute. Du fait des opérations de morphologie mathématique et des seuils du nombre de population (en nombre de pixels) lors de la segmentation, certains arbres isolés sont omis des masques. Mais ces cas restent marginaux.



FIG. 6.14 – Exemple d'arbres isolés exclus du masque final de la végétation haute.

Il serait bien sûr possible d'envisager une autre méthode. Par exemple, la littérature propose plusieurs méthodes de détection des nuages, comme par exemple [Béréziat, 1999]. Puis, grâce aux informations sur les positions du satellite et du soleil, il suffirait de projeter le nuage segmenter au sol pour obtenir la région correspondante à son ombre portée.

6.3 Conclusion

La mise en place d'une méthode totalement automatique de dénombrement de cocotiers est particulièrement difficile dans notre cas. La première difficulté vient du manque de la bande proche infra-rouge qui permet de calculer un indice de végétation (le NDVI). Ce problème est résolu au prix de l'intervention de l'utilisateur qui choisit les zones d'apprentissages pour la détection des zones végétales par la classification bayésienne. Du fait de la nature du sol des atolls (présence de sable), il est nécessaire de réaliser cette étape au mieux pour éviter des erreurs de détection de houppiers sur du sable (recherche de maxima locaux des intensités de niveaux de gris). L'obtention d'images dans le proche infra-rouge est utile pour limiter l'intervention de l'utilisateur en ce qui concerne la partie pré-traitement (génération des masques).

La mission terrain de 2006 a permis de rendre compte de trois différentes classes (strates) de cocotiers présentes dans les cocoteraies. Nous avons montré que certains arbres étaient partiellement ou entièrement masqué par la palmeraie d'un arbre voisin, et de ce fait, une détection de ce type d'arbre n'est pas possible. Une autre difficulté rencontrée concernent les arbres trop proches, c'est-à-dire ceux de appartenant à la strate $S2$ ou deux arbres de la strate $S1$ avec un chevauchement important de leurs palmes. Nous nous trouvons dans une situation d'agglomération d'arbre et dans ce cas nous avons eu des difficultés pour discriminer ces zones. La figure 6.15 présente un exemple d'agglomération non séparée après la segmentation et quelques types d'agglomérations d'arbres possibles.

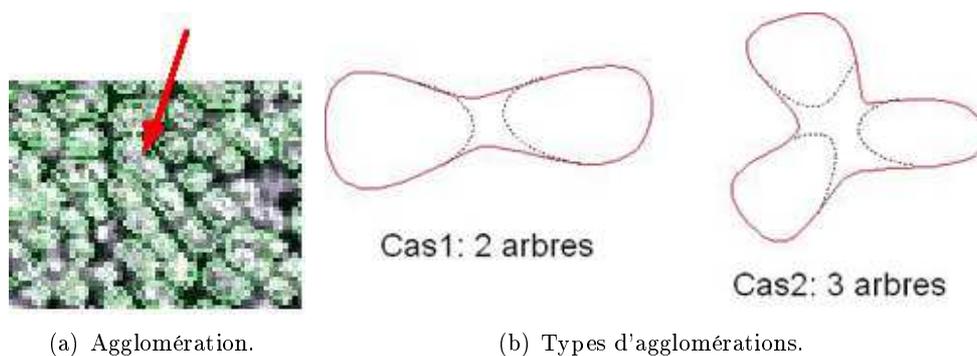


FIG. 6.15 – Exemples d'erreurs de segmentation et type d'agglomération.

Nos images datent de 2003 et lors de la mission terrain sur Tikehau, effectuée en 2006, nous avons pu nous rendre compte qu'une grande surface de la cocoteraie a été défrichée pour un projet de scierie afin de permettre l'exploitation du bois des cocotiers séniles. Il se peut qu'en l'espace de trois années, des arbres juvéniles à l'époque (non visible sur les images) aient suffisamment poussé pour atteindre une hauteur et une surface de sa palmeraie assez grande pour être visible sur les images satellites. La commande d'images plus récentes est une nécessité : elle permettra de comparer les résultats obtenus et ainsi obtenir des résultats plus proche de la réalité.

Une étude plus avancée sur l'analyse spatiale de la disposition des centres des houppiers donnerait une indication sur le type de cocoteraie (sauvage, artificielle ou abandonnée) qui serait très utile pour l'exploitation des cocoteraies. Deux pistes sont explorées dans la dernière partie de cette thèse. La première approche consiste à caractériser les différents types de plantation grâce au informations spectrales et de texture. La seconde approche est d'utiliser les outils d'analyse spatiale pour traiter les coordonnées des arbres afin de caractériser les types de plantation à l'aide de processus ponctuel de poisson.

Détection des ombres et estimation de la hauteur des arbres

Le problème de la segmentation des ombres est traité dans ce chapitre. Deux méthodes de l'état de l'art qui ont l'avantage d'être simples et performantes sont présentées. Une méthode géométrique d'estimation de la hauteur des cocotiers, à partir de leur ombre portée, est proposée.

7.1 Détection des ombres

Les ombres se produisent lorsqu'un objet se trouve devant une source d'illumination. On distingue deux catégories :

- l'ombre portée : c'est le résultat de la projection de l'objet au sol dans la direction de la source de lumière,
- la zone ombragée : la région de l'objet qui ne se trouve pas éclairée.

Les méthodes de détection des ombres se déclinent en deux classes :

- Méthodes basées sur des modèles [Thirion, 1992, Koller et al., 1993, Bejanin et al., 1994]. Ces méthodes détectent les ombres en se basant sur une connaissance *a priori* de la géométrie de la scène, des objets et de l'illumination. Elles sont limitées car toutes les images ne disposent pas de ces informations.
- Méthodes basées sur des propriétés optiques. Ces méthodes identifient les ombres en utilisant la luminosité ou la couleur de l'ombre [Salvador et al., 2004, Tsai, 2003, Tsai, 2004, Polidorio et al., 2003, Huang et al., 2004].

Dans cette étude, nous étudions la seconde approche. Dans [Salvador et al., 2004], les auteurs proposent de détecter les ombres en utilisant, en plus de l'espace RGB normalisé et l'espace HSI, des espaces de couleurs invariants aux variations photométriques [Gevers and Smeulders, 1999]. Nous utilisons et présentons dans la section suivante les méthodes de Polidorio [Polidorio et al., 2003] et de Huang [Huang et al., 2004] pour segmenter les ombres dans les images Ikonos.

7.1.1 La méthode de Polidorio

Deux propriétés sont utilisées pour réaliser la segmentation des zones d'ombres. La première est que les zones d'ombres ont une faible luminance, ceci parce qu'elles ne reçoivent pas directement le flux de lumière à cause des obstacles se trouvant entre elles et la source de lumière. La seconde prend en considération le modèle de dispersion atmosphérique de Rayleigh assumant que les zones d'ombres sont fortement saturées en bleu/violet par rapport aux autres longueurs d'ondes du spectre visible. Dans ces conditions, les zones d'ombres doivent donc vérifier deux propriétés :

- forte saturation du bleu/violet,
- luminance faible.

En changeant d'espace de couleurs, c'est-à-dire en passant de l'espace RGB à l'espace HSI, il est aisé de caractériser ces propriétés. La figure 7.1 montre les différentes composantes obtenues lors du changement de l'espace de couleurs RGB en HSI.

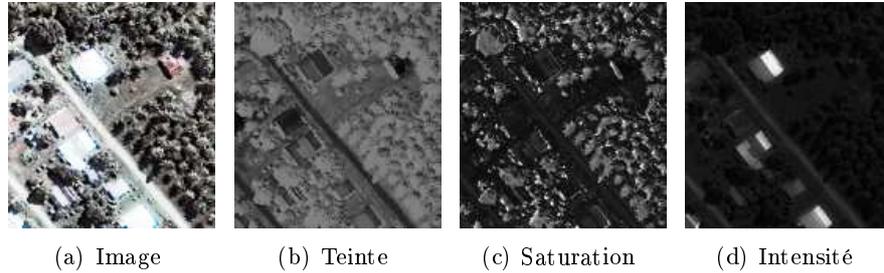


FIG. 7.1 – Transformation de l'image RGB en HSI.

La méthode consiste à produire une carte des ombres, notée SDW , par seuillage de la différence entre l'image d'intensité I et l'image de saturation S .

$$SDW_{ij,k} = \begin{cases} 1 & , \text{ si } I_{ij} - S_{ij} \leq K \\ 0 & , \text{ sinon} \end{cases} \quad (7.1)$$

$$K = \begin{cases} 0 & , \text{ pour les systèmes aéroportés} \\ -0.2 & , \text{ pour les satellites} \end{cases} \quad (7.2)$$

Le seuil K est choisi par calibration. Cette valeur est dépendante de la distance entre le capteur et l'objet. L'intensité de la dispersion atmosphérique de Rayleigh est directement proportionnelle à la distance entre la scène et le capteur (une faible distance implique que la concentration de particules en suspension dans l'air est moins importante ce qui en résulte une faible dispersion de la lumière du soleil).

La figure 7.2 présente le résultat obtenu sur l'image de la figure 7.1(a). Le contour des ombres est affiché en jaune et en superposition à l'image originale. Nous constatons que la méthode donne de bons résultats. Elle permet de détecter les ombres portées au sol des maisons (voir la maison au toit rouge) et des arbres sur différents types de sol. En effet, l'ombre des arbres est aussi bien détectée sur de l'herbe que sur du sable.



FIG. 7.2 – Résultat de la segmentation des zones d'ombres avec la méthode de Polidorio.

7.1.2 La méthode de Huang

Cette méthode segmente les ombres à partir des composantes bleue (B), verte (G) et la teinte (en anglais *Hue*). Il convient d'abord de calculer la teinte :

$$H = \arctan \left(\frac{\sqrt{3}(G - B)}{(R - G) + (R - B)} \right) \quad (7.3)$$

Les zones d'ombres possèdent des valeurs élevées dans la teinte (composante H), des valeurs faibles dans le bleu et des valeurs faibles dans la différence du vert et du bleu. Une technique de seuillage est utilisée pour détecter les zones d'ombres. L'algorithme est le suivant :

Soit $I(i, j)$ une image couleur RGB.

- 1 Transformer l'image *RGB* en image *HSI*. Un seuil optimal τ_1 est calculé ([Otsu, 1979] par exemple) pour trouver dans la composante H des régions d'ombres candidates :

$$S_c(i, j) = \mathbb{1}_{H(i, j) > \tau_1} \quad (7.4)$$

S_c peut contenir les ombres mais aussi les objets bleus et verts.

- 2 Pour éliminer les objets bleus de S_c , un autre seuil, τ_2 , est calculé avec la méthode de Otsu et utilisé pour segmenter la bande B tel que :

$$S_{cc}(i, j) = S_c(i, j) \times \mathbb{1}_{B(i, j) < \tau_2} \quad (7.5)$$

S_{cc} ne contient plus que les objets verts à éliminer.

- 3 Pour éliminer les objets verts de S_{cc} , un troisième seuil τ_3 est calculé avec la méthode de Otsu et utilisé tel que :

$$S_{ccc}(i, j) = S_{cc}(i, j) \times \mathbb{1}_{G(i, j) - B(i, j) < \tau_3} \quad (7.6)$$

S_{ccc} ne contient plus que les objets verts à éliminer.

- 4 Utilisation d'une d'ouverture morphologique pour éliminer les points isolés et de fermeture morphologique remplir les petites régions.

Pour illustrer cette méthode, nous travaillons sur l'image de la figure 7.3.



FIG. 7.3 – Image de test pour la méthode.

La figure 7.4 montre les étapes de la segmentation des ombres de la méthode présentée ci-dessus. Sur la première ligne, il est présenté différents histogrammes avec les valeurs de seuils estimé par la méthode d'Otsu. En dessous, ce sont les images seuillées qui sont présentées. La figure 7.5 présente le résultat de la segmentation des ombres obtenues par la méthode Huang. Si les résultats obtenus sont satisfaisant, ils ne sont pas aussi performants que ceux obtenus avec la méthode de Polidorio. Par la suite, nous utilisons cette dernière pour segmenter les ombres puis estimer la hauteur des arbres.

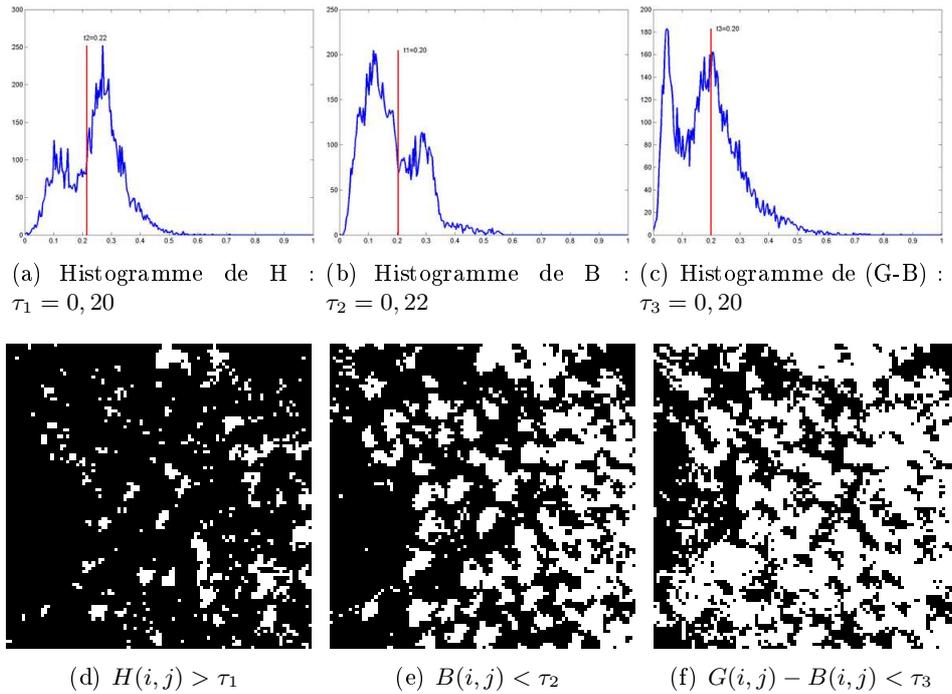


FIG. 7.4 – Les différentes étapes de la méthode de Huang.

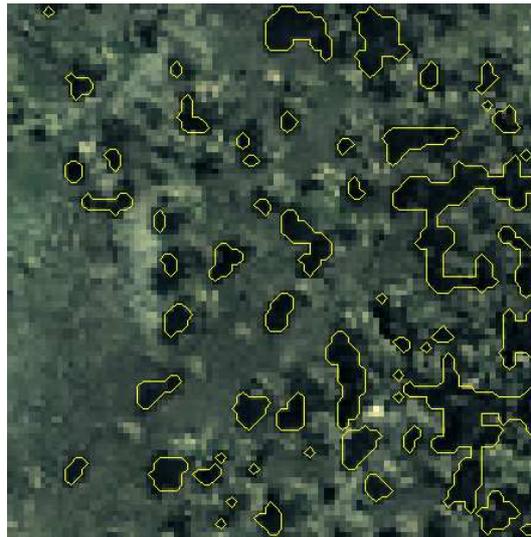


FIG. 7.5 – Résultat de la segmentation des ombres grâce à la méthode de Huang.

7.2 Estimation de la hauteur à partir de l'ombre portée

Dans cette section, nous nous intéressons à une méthode géométrique pour calculer la hauteur d'un arbre à partir de la longueur de l'ombre portée visible par le satellite, la position du satellite ainsi que celle du soleil. Ensuite, nous appliquons cette méthode pour estimer la hauteur des cocotiers.

7.2.1 Méthode d'estimation des hauteurs

Dans [Shettigara and Sumerling, 1998] est décrite une méthode permettant d'estimer la hauteur des objets en fonction de la longueur des ombres portées au sol, la position du soleil et celle du capteur. L'article présente l'application de cette méthode sur des images SPOT.

avec :

- θ_{sun} : l'angle d'élévation du soleil par rapport au sol,

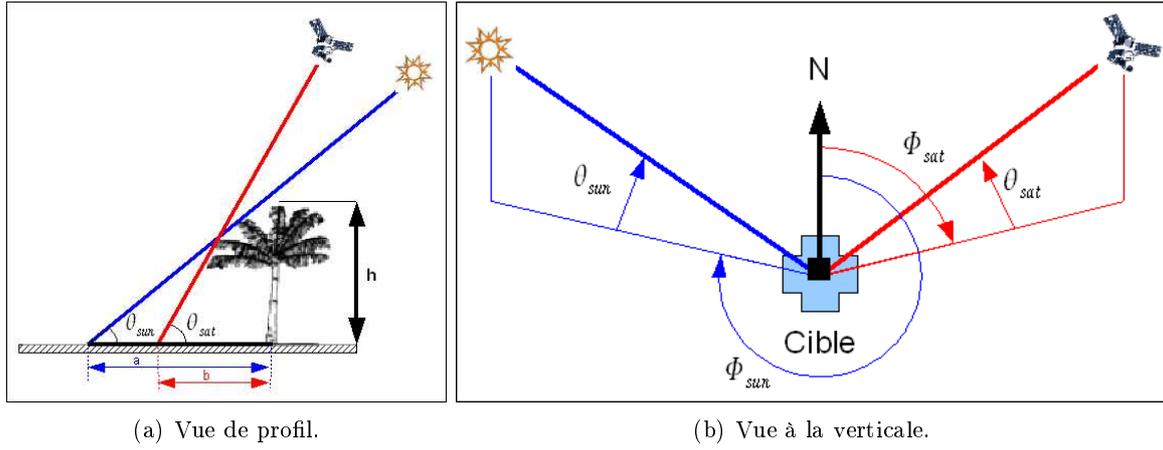


FIG. 7.6 – Configuration de la position du soleil, du capteur et de l'arbre.

- θ_{sat} : l'angle d'élevation du satellite (*Nominal Collection Elevation* dans le fichier METADATA)
- h : la hauteur de l'arbre,
- a : longueur totale de l'ombre portée,
- b : longueur de l'ombre non visible du satellite.

Comme le montre la figure 7.6(a), l'ombre visible par le satellite est dépendante de sa position par rapport au soleil et celles des arbres. Si le satellite se trouve à l'opposé du soleil, la zone d'ombre a est alors entièrement visible. Si le satellite se trouve du même côté que le soleil, alors la zone d'ombre a est alors obstruée par l'arbre et la longueur visible est alors de $a - b$, avec :

- θ_{sun} : l'angle d'élevation du soleil par rapport au sol,
- θ_{sat} : l'angle d'élevation du satellite (*Nominal collection elevation* dans le fichier METADATA),
- ϕ_{sun} : azimuth du soleil (angle par rapport au Nord),
- ϕ_{sat} : azimuth du satellite (*Nominal Collection Azimuth* dans le fichier METADATA) .

Grâce à ϕ_{sun} , il est possible de calculer la direction de la source de lumière et ainsi avoir une direction dans laquelle chercher des zones d'ombres. D'après la figure 7.6(a), la longueur de l'ombre le long de la direction donnée par le soleil est :

$$a = \frac{h}{\tan(\theta_{sun})} \quad (7.7)$$

La longueur de l'ombre obstruée par un objet vue depuis le satellite est donnée par la relation suivante :

$$b = \frac{h}{\tan(\theta_{sat})} \quad (7.8)$$

La composante de l'ombre des objets le long de la normale d'une ligne d'arbres est :

$$S_{sun} = a \cos(\phi_{su}) \quad (7.9)$$

Celle de l'ombre obstruée par des objets le long de la normale d'une ligne d'arbres est :

$$S_{sat} = b \cos(\phi_{sa}) \quad (7.10)$$

avec

$$\phi_{su} = \phi_{sun} + \frac{\pi}{2} - \phi_t \quad (7.11)$$

$$\phi_{sa} = \phi_{sat} + \frac{\pi}{2} - \phi_t \quad (7.12)$$

où ϕ_t est l'azimut d'une ligne d'arbres dont on souhaite estimer la hauteur.

Étant donné que la longueur de l'ombre vue depuis le satellite et le long de la normale de ligne d'arbres est $S = S_{sun} - S_{sat}$, on peut en déduire la hauteur de la ligne d'arbres avec la relation suivante :

$$h = \frac{s}{\left[\frac{\cos(\phi_{su})}{\tan(\theta_{sun})} \right] - \left[\frac{\cos(\phi_{sa})}{\tan(\theta_{sat})} \right]} \quad (7.13)$$

Le tableau 7.1 présente la structure d'un fichier de méta données fourni avec les images, les différentes valeurs nécessaires pour l'estimation de la hauteur sont données :

Nominal Collection Azimuth	281.7026 degrés
Nominal Collection Elevation	61.44155 degrés
Sun Angle Azimuth	28.1930 degrés
Sun Angle Elevation	47.11778 degrés

TAB. 7.1 – Quelques valeurs du fichier de méta données.

7.3 Application à l'estimation des hauteurs des cocotiers

Nous appliquons cette méthode d'estimation des hauteurs aux cocotiers dans les images IKONOS. Prenons le résultat de la détection des ombres illustré par la figure 7.5 dont la carte des ombres est présentée dans la figure 7.7. L'illustration de la méthode d'estimation des hauteurs à partir des ombres

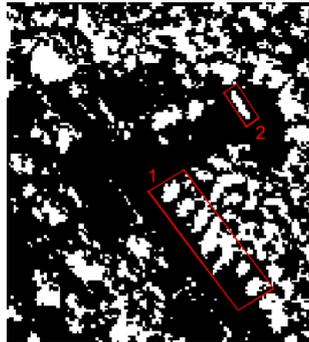
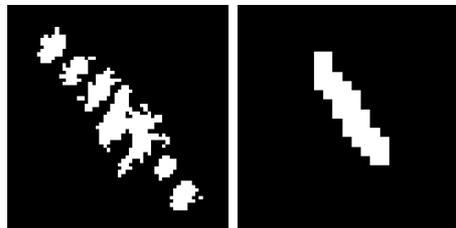


FIG. 7.7 – Carte des ombres.

portées est réalisée sur deux lignes d'ombres dont un zoom est présenté dans la figure 7.8.



(a) Ligne d'ombre 1. (b) Ligne d'ombre 2.

FIG. 7.8 – Choix des deux lignes d'ombres pour l'estimation de la hauteur.

À partir de ces images, il est nécessaire de calculer ϕ_t l'azimut de la ligne d'arbre. Celle-ci est réalisée par régression linéaire. Les résultats de cette estimation de l'orientation des lignes d'ombres sont présentés dans la figure 7.9. La ligne d'ombre 1 présente un groupement d'arbres qui sont numérotés (voir figure 7.9(a)). La ligne d'ombre 2 est l'ombre portée au sol d'un bâti qui nous sert à valider l'estimation de la hauteur. En effet, si l'estimation de la hauteur de ce bâti donne un résultat satisfaisant, nous pouvons en déduire que ceux obtenus pour les arbres le sont aussi.

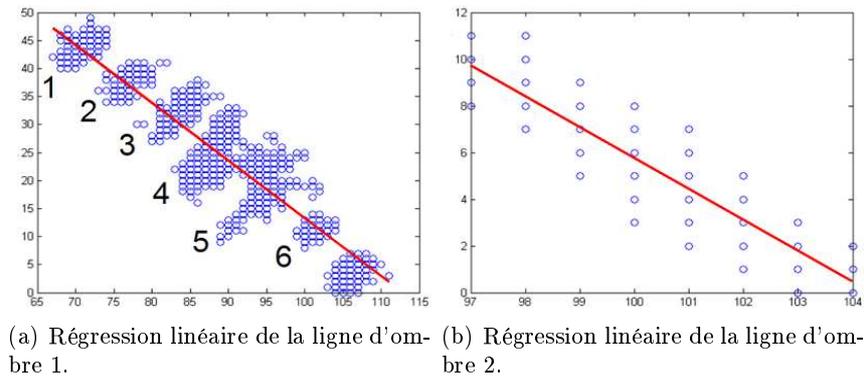


FIG. 7.9 – Régression linéaire des lignes d'ombres pour estimer l'azimut.

La longueur S de l'ombre visible par le satellite est calculée dans une direction perpendiculaire à l'azimut de la ligne d'arbres. Grâce à l'équation 7.13, nous pouvons estimer la hauteur de chacun de ces arbres ainsi que celle du bâti. Les résultats sont présentés dans le tableau 7.2. Nous observons que le bâti a une hauteur estimée de 2,39 m, ce qui reste tout à fait acceptable pour le type de bâti présents sur l'atoll et dans les cocoteraies (constructions très basses qui servent de lieu de repos ou de restauration lors de la récolte du *Coprah*). De plus, les hauteurs estimées des cocotiers restent cohérentes avec la hauteur maximale que peut avoir un cocotier "Grand Local" qui est de 25 mètres.

Objet	Hauteur estimée
arbre 1	7,4 m
arbre 2	6,47 m
arbre 3	8,33 m
arbre 4	13,88 m
arbre 5	12,95 m
arbre 6	4,63 m
bâti	2,39 m

TAB. 7.2 – Résultats de l'estimation des hauteurs.

Nous venons de voir qu'il est possible, à partir des ombres détectées sur les images Ikonos, d'estimer la hauteur des cocotiers. Cependant, cette méthode connaît des limitations :

- la longueur de l'ombre doit être de 1 pixel minimum et continue,
- on suppose que le sol est plat (ce qui est le cas pour les atolls),
- la hauteur n'est pas estimable si l'ombre n'est pas projeté au sol mais sur un autre objet de l'image.

Afin d'estimer la hauteur d'un ensemble d'arbre, il serait judicieux de considérer que tous les cocotiers d'une plantation artificielle ont été plantés au même moment (ou dans un laps de temps très court). Il devient alors évident que ces arbres ont une croissance similaire et donc une hauteur proche. De plus, la hauteur d'un cocotier permet d'estimer son âge et donc son potentiel de production de noix de coco.

Troisième partie

Typologie de la cocoteraie

État de l'art sur la classification

Dans ce chapitre, il est présenté un état de l'art sur la classification en traitement de l'image. Les algorithmes de classification supervisés les plus couramment utilisés en imagerie satellitaires sont présentés.

8.1 Définition de la classification

La classification en traitement de l'image consiste à attribuer à chaque pixels de l'image une classe. La classe est calculée, on dit décidée, à partir d'une fonction (ou règle) de décision appliquée aux valeurs radiométriques de l'image. La classification permet de créer une carte d'occupation des sols, de caractériser des zones végétales, ...

L'attribution de la classe ω au pixel x est établie par une **fonction discriminante** $g_\omega(x)$ qui évalue la vraisemblance de l'appartenance de x à la classe ω . Ainsi la **règle de décision** pour attribuer la classe ω au pixel x est de retenir la classe pour laquelle la fonction discriminante est la plus grande pour ce pixel :

$$x \in \omega, \text{ si } g_\omega(x) > g_{\omega'}(x) \quad \forall \omega' \neq \omega \quad (8.1)$$

On distingue deux types de classification :

- les classifications non supervisées (ou automatiques) qui classent les pixels sur des critères statistiques et sans avoir recours à des échantillons d'entraînement. Ces méthodes attribuent aux régions ayant une caractéristique spectrale similaire la même classe.

Les principaux algorithmes de classification non supervisés sont : le regroupement par k -moyennes (*K-Means*) [McQueen, 1967, Forgy, 1965] et sa variante, le regroupement par *ISODATA* (*Iterative Self-Organizing Data Analysis*) [Ball and Hall, 1965].

- les classifications supervisées qui se basent sur une connaissance *a priori* de l'occupation des sols. L'utilisateur doit définir des zones d'apprentissage pour entraîner l'algorithme de classification à reconnaître les différentes classes en fonction de leurs signatures spectrales. Un ensemble d'apprentissage est défini par un ensemble de couple $\{X_i, \omega_i\}_{1 \leq i \leq n}$ représentant n réalisations indépendantes d'une variable aléatoire (X, ω) de loi inconnue P . L'objectif de la classification supervisée est de prédire ω_{n+1} à partir de X_{n+1} et $\{X_i, \omega_i\}_{1 \leq i \leq n}$ telle que (X_{n+1}, ω_{n+1}) soit une réalisation de (X, ω) et indépendante de $\{X_i, \omega_i\}_{1 \leq i \leq n}$. Une fois cette connaissance acquise par l'algorithme, il ne lui reste plus qu'à extrapoler à l'ensemble des pixels de l'image. Ce sont ces méthodes qui sont les plus utilisées dans le traitement des images satellites.

Dans notre étude, nous allons utiliser uniquement des algorithmes de classification supervisée car elles sont plus performantes que les non supervisées. Les sections suivantes présentent les algorithmes les plus couramment utilisés dans la classification supervisée, des méthodes les plus simples aux méthodes les plus élaborées. Nous passerons donc en revue sur la classification des k -plus proches voisins, la classification par Distance Minimum, la classification par Distance de Mahalanobis, la classification

Bayésienne par Maximum de Vraisemblance et ainsi les Séparateurs à Vaste Marge (*Support Vector Machine* ou SVM). Ces classifieurs sont comparés entre eux en analysant leurs performances pour deux tâches de classification :

1. classifier l'occupation des sols. Ceci nous permet d'estimer les surfaces des cocoteraies (tout type confondu),
2. classifier les types de cocoteraies selon leur organisation spatiale.

8.2 k -plus proches voisins

L'algorithme de classification des k -plus proches voisins est particulièrement simple dans son principe. Il suppose que des pixels proches les uns des autres dans un espace de description sont susceptibles d'appartenir à la même classe. Un pixel est étiqueté en examinant l'ensemble des pixels à classifier et en lui attribuant l'étiquette de la classe la plus représentée parmi un nombre prédéfini de plus proches voisins. La comparaison nécessite essentiellement le calcul de distance entre la couleur du pixel à étiqueter et celle des autres pixels.

On se donne M classes notées $1, \dots, M$. On suppose qu'il y ait N_i voisins étiquetés comme appartenant à la classe i parmi les k plus proches voisins d'un pixel x . On note ω_x la classe du pixel x . Remarquons que $\sum_{i=1}^M N_i = k$. La fonction discriminante pour la $i^{\text{ème}}$ classe est définie telle que $g_i(x) = N_i$ avec la règle de décision :

$$\omega_x = i \text{ si } g_i(x) > g_j(x) \forall i \neq j \quad (8.2)$$

Mais cette règle de décision ne prend pas du tout en compte la distance spectrale entre le pixel à étiqueter et ses plus proches voisins. Une autre règle de décision plus performante peut être définie par :

$$g_i(x) = \frac{\sum_{j=1}^{N_i} \frac{1}{d(x, x_j^i)}}{\sum_{i=1}^M \sum_{j=1}^{N_i} \frac{1}{d(x, x_j^i)}} \quad (8.3)$$

$d(x, x_j^i)$ la distance spectrale (le plus souvent la distance euclidienne dans l'espace des couleurs) entre le pixel x à étiqueter et ses voisins x_j^i (le $j^{\text{ème}}$ des N_i pixels de la classe i).

8.3 Classification par parallélépipèdes

La classification par parallélépipèdes est basée sur l'examen individuel des histogrammes des composants spectraux des données d'apprentissage. Les bornes inférieures et supérieures de l'histogramme sont identifiées et utilisées pour décrire l'intervalle de valeur des niveaux de gris de chaque bande pour une classe donnée. L'ensemble des intervalles forment alors une boîte multidimensionnelle ou parallélépipèdes. Lors de la classification, les pixels qui seront trouvés dans un tel parallélépipède seront alors étiquetés comme appartenant à la classe correspondante. La figure 8.3 montre de tels parallélépipèdes pour chacune des classes représentées par une couleur différente.

Cette méthode de classification est simple et rapide mais possède quelques inconvénients :

- il peut exister de grands espaces vides entre les parallélépipèdes et les pixels de ces régions ne seront pas classifiés,
- les probabilités *a priori* d'appartenance à une classe ne sont pas prises en compte,
- lorsque les données sont corrélées, il peut alors exister des chevauchement des parallélépipèdes, et dans ce cas les données ne seront pas séparables.

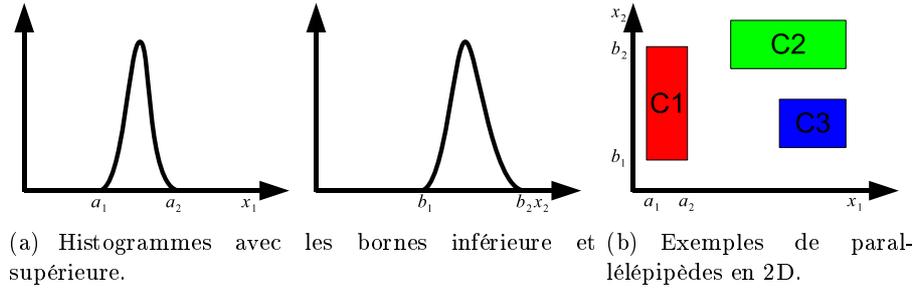


FIG. 8.1 – Principe de la classification par parallélépipèdes.

8.4 Classification par distance minimum

Lorsque les données d'apprentissage pour chaque classe sont de faible taille ou peu représentative, un classifieur ne nécessitant pas l'information sur la dispersion (c'est-à-dire covariance) mais uniquement la position moyenne des classes spectrales peut suffire pour donner de bons résultats. Les données d'apprentissage sont utilisés pour déterminer la moyenne des classes et un pixel sera étiqueté dans la classe ayant la moyenne la plus proche.

Soit m_i , $i = 1, \dots, M$ les moyennes des M classes déterminées à partir des données d'apprentissage, et x l'information multi-spectrale d'un pixel à classifier. On calcule l'ensemble des distances euclidiennes élevées au carré du pixel à classifier par rapport aux moyennes des classes :

$$\begin{aligned} d(x, m_i)^2 &= (x - m_i)^T (x - m_i) \\ &= x^T x - 2m_i^T x + m_i^T m_i \end{aligned} \quad (8.4)$$

La règle de classification utilisée devient alors :

$$\omega_x = i \text{ si } d(x, m_i)^2 < d(x, m_j)^2 \quad \forall i \neq j$$

étant donné que le terme $x^T x$ est commun à tous les $d(x, m_i)^2$ et en inversant les signes, la classification est réalisé en se basant sur la règle suivante :

$$\omega_x = i \text{ si } g_i(x) > g_j(x) \quad \forall i \neq j$$

avec $g_i(x) = 2m_i^T x - m_i^T m_i$.

Dans le cas où un nouveau pixel à classifier est équidistant dans l'espace de représentation des classes, ce dernier n'est pas classifier.

8.5 Classification par distance de Mahalanobis

La classification par distance de Mahalanobis est une amélioration de celle par distance minimum en calculant les distances interclasses dans la base où les valeurs sont les plus décorréelées entre elles. Ainsi, nous obtenons une meilleur discrimination interclasse.

En considérant le fonction discriminante de la classification par Maximum de Vraisemblance dans le cas où les probabilités *a priori* sont supposées égales, on a

$$g_i(x) = -\ln(|\Sigma_i|) - (x - m_i)^T \Sigma_i^{-1} (x - m_i)$$

avec Σ_i la covariance de la classe i .

En inversant le signe de la fonction ci-dessus, celle-ci peut être considérée comme une mesure de distance élevée au carré. La mesure de distance s'exprime alors :

$$d(x, m_i)^2 = \ln(|\Sigma_i|) + (x - m_i)^T \Sigma_i^{-1} (x - m_i) \quad (8.5)$$

Si on considère que les covariance des classes sont égales, c'est-à-dire que $\Sigma_i = \Sigma \quad \forall i$, la mesure de distance peut être réécrite telle que :

$$d(x, m_i)^2 = (x - m_i)^T \Sigma^{-1} (x - m_i) \quad (8.6)$$

Ce type de classifieur est connu sous le nom de classifieur par distance de Mahalanobis car la distance de Mahalanobis est la racine carrée de l'équation (8.6).

8.6 Classification Bayésienne

Chaque pixel i de l'image est associé à une variable aléatoire X_i qui décrit sa couleur et une seconde variable aléatoire W_i qui décrit sa classe. Les couples (X_i, W_i) sont indépendants et identiquement distribués. On a :

$$X_i \in \{1, \dots, N\}$$

$$\text{et } W_i \in \{1, \dots, M\}$$

Définition 8.1 *La classification consiste à considérer les données et classes comme des processus stochastiques dont on connaît la loi. La fonction discriminante est alors calculée à partir des lois conditionnelles des données à classer.*

Pour déterminer la classe à laquelle le pixel i appartient, on choisit l'étiquette ω_i qui maximise la probabilité conditionnelle $P(W_i|X_i = x_i)$, c'est-à-dire la réalisation ω_i de W_i la plus probable sachant que le pixel i a la couleur x_i . Cette règle de décision s'écrit :

$$W_i = \omega_i \text{ si } P(W_i = \omega_i|X_i = x_i) \geq P(W_i = \omega|X_i = x_i) \quad \forall \omega \neq \omega_i \quad (8.7)$$

Approche Bayésienne

Les probabilités $P(W_i = \omega_i|X_i = x_i)$ ne sont pas connues mais grâce au théorème de Bayes, nous pouvons calculer cette loi. Définissons d'abord les lois de probabilités suivantes :

$P(X_i = x_i|W_i = \omega_i)$: la probabilité que le pixel i a la couleur x_i sachant qu'il appartient à la classe ω_i . On notera cette probabilité, par abus de notation et quand il n'y a pas d'ambiguïté, $P(x_i|\omega_i)$.

$P(W_i = \omega_i|X_i = x_i)$: la probabilité que le pixel i appartienne à la classe ω_i sachant qu'il a la couleur x_i . On notera $P(W_i = \omega_i|X_i = x_i) = P(\omega_i|x_i)$ par abus de notation.

$P(W_i = \omega_i)$: la probabilité que le pixel i appartienne à la classe ω_i . On notera cette probabilité $P(\omega_i)$.

$P(X_i = x_i)$: la probabilité que le pixel i ait la couleur x , on notera cette probabilité $P(x_i)$.

Ainsi, en appliquant le théorème de Bayes, la probabilité conditionnelle de W_i sachant X_i s'écrit :

$$P(\omega_i|x_i) = \frac{P(x_i|\omega_i)P(\omega_i)}{P(x_i)} \quad (8.8)$$

Pour rendre possible le calcul de $P(\omega_i|x_i)$, on doit faire une hypothèse de distribution pour la loi conditionnelle $P(x_i|\omega_i)$. Cette loi décrit la répartition des niveaux de gris de l'image sous hypothèse de classe ω_i et nous la supposons gaussienne car elle rend bien compte d'une répartition homogène pour chaque classe comme constaté dans nos données. Ainsi, pour chaque classe, il faut donc estimer les paramètres de la loi gaussienne (moyenne et covariance).

L'apprentissage et règle de décision

La loi conditionnelle $P(x_i|\omega_i)$ est supposée gaussienne. Sa densité s'écrit donc :

$$p(x_i|\omega_i) = \frac{1}{(2\pi^N |\Sigma_{\omega_i}|)^{-1}} \exp\left(-\frac{1}{2}(x_i - \mu_{\omega_i})^T \Sigma_{\omega_i}^{-1} (x_i - \mu_{\omega_i})\right) \quad (8.9)$$

La moyenne et la covariance $(\mu_{\omega}, \Sigma_{\omega})$ de la classe ω sont donnés par les estimateurs du maximum de vraisemblance calculés sur les ensembles d'apprentissages :

$$\hat{\mu}_{\omega} = \frac{1}{N_{\omega}} \sum_i x_i \mathbb{1}_{\omega_i=\omega}, \quad N_{\omega} = \sum_i \mathbb{1}_{\omega_i=\omega}$$

$$\hat{\Sigma}_{\omega} = \frac{1}{N_{\omega} - 1} \sum_i (x_i - \hat{\mu}_{\omega})^T (x_i - \hat{\mu}_{\omega}) \mathbb{1}_{\omega_i=\omega}$$

Les lois $P(x_i)$ et $P(\omega_i)$ peuvent être estimées à partir respectivement de l'histogramme des couleurs et de l'histogramme des étiquettes calculés sur les ensembles d'apprentissage. En pratique, on ne calcule pas $P(x)$ car cette donnée se déduit de la loi conditionnelle des x_i sachant ω_i et de la loi de ω_i :

$$\begin{aligned} P(X_i = x_i) &= \sum_l P(X_i = x_i, W_i = l) \\ &= \sum_l P(X_i = x_i | W_i = l) P(W_i = l) \end{aligned} \quad (8.10)$$

En injectant dans l'équation (8.8), on obtient :

$$P(\omega_i | x_i) = \frac{P(x_i | \omega_i) P(\omega_i)}{\sum_l P(X_i = x_i | W_i = l) P(W_i = l)} \quad (8.11)$$

La règle de décision est de choisir la classe ω_i qui maximise $P(\omega_i | x_i)$, ainsi on a :

$$\begin{aligned} &P(\omega_i | x_i) \geq P(\omega | x_i) \quad \forall \omega \neq \omega_i \\ \Leftrightarrow &\frac{P(x_i | \omega_i) P(\omega_i)}{P(x_i)} \geq \frac{P(x_i | \omega) P(\omega)}{P(x_i)} \quad \forall \omega \neq \omega_i \\ \Leftrightarrow &P(x_i | \omega_i) P(\omega_i) \geq p(x_i | \omega) P(\omega) \end{aligned}$$

Le calcul des $P(x_i)$ n'est donc pas utile. Enfin, on remplace la loi conditionnelle par la loi gaussienne et on applique le logarithme à l'inéquation (8.6), on obtient :

$$-\frac{1}{2}(x_i - \mu_{\omega_i})^T \Sigma_{\omega_i}^{-1} (x_i - \mu_{\omega_i}) - \frac{1}{2} \log |\Sigma_{\omega_i}^{-1}| + \log P(\omega_i) \geq -\frac{1}{2}(x_i - \mu_{\omega})^T \Sigma_{\omega}^{-1} (x_i - \mu_{\omega}) - \frac{1}{2} \log |\Sigma_{\omega}^{-1}| + \log P(\omega) \quad (8.12)$$

On retrouve de part et d'autre la distance de Mahalanobis à laquelle est additionné un biais b qui dépend des lois *a posteriori* :

$$\begin{cases} d_{\Sigma_{\omega}}(x_i, \mu_{\omega_i}) \geq d_{\Sigma_{\omega}}(x_i, \mu_{\omega}) + b \\ b = \log \frac{P(\omega)}{P(\omega_i)} \end{cases} \quad (8.13)$$

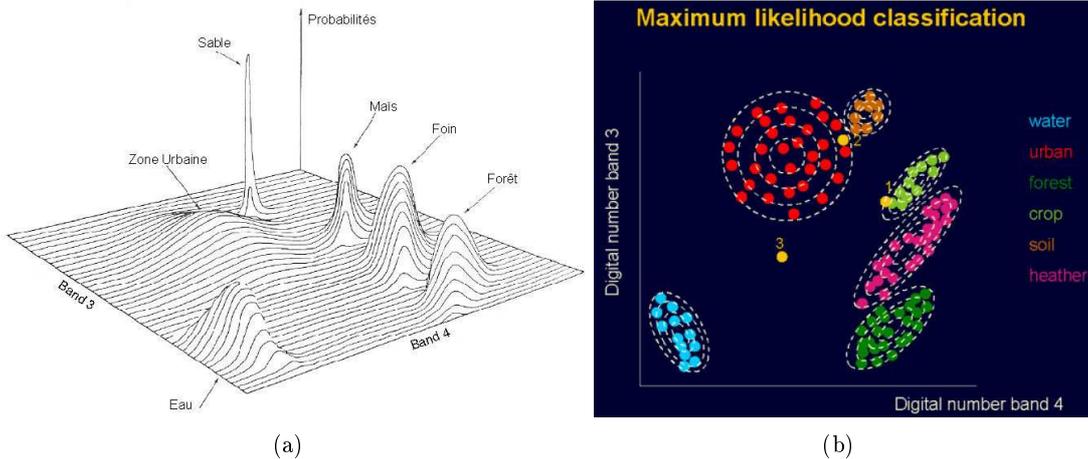


FIG. 8.2 – Illustration de la distribution des densité de probabilité.

8.7 Séparateurs à Vaste Marge

8.7.1 Introduction

Les *Support Vector Machines* ou *Séparateur à Vaste Marge* (SVM) forment une classe d'algorithmes d'apprentissage initialement définis pour la discrimination binaire. Ils sont basés sur la

recherche d'*hyperplans de marge optimale* qui classent ou séparent en deux parties des données tout en étant le plus éloignés possible de toutes les observations. Ce sont donc des classificateurs binaires. Ces méthodes ont été introduites par [Cortes and Vapnik, 1995]. Le principe est de ramener le problème de la discrimination à la recherche d'un hyperplan optimal :

- dans le cas où les données sont linéairement séparables, le problème revient à chercher un hyperplan optimal dont la marge avec les données est la plus grande,
- dans le cas où les données ne sont pas linéairement séparables, il faut cette fois-ci minimiser les erreurs de classification,
- et enfin, dans le cas où les données sont hautement non linéaires, l'idée est de projeter les données dans un espace de dimension supérieure dans lequel le problème redevient linéairement séparable.

Les SVM nécessitent un jeu de données d'apprentissage pour estimer les paramètres du modèle. Étant donné un ensemble d'apprentissage $\mathcal{A}_n = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ où x_i représente la donnée (ou observation) et y_i la décision associée appartenant à $\{-1, +1\}$, nous cherchons l'hyperplan $x \cdot \omega + b = 0$ qui est situé à une distance maximale des x_i .

8.7.2 Le plan séparateur et la marge

a) Équation d'un plan dans l'espace

Commençons par rappeler l'équation d'un plan. Un point x appartient au plan P si on a la relation suivante :

$$x \in P \Leftrightarrow x \cdot \omega + b = 0 \quad (8.14)$$

où

- ω est un vecteur normalisé et orthogonal au plan P ,
- $b = -x_M \cdot \omega$ avec x_M un point du plan P .

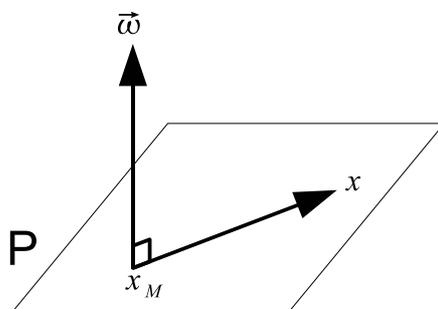


FIG. 8.3 – Position d'un point x par rapport à un plan P .

b) Distance d'un point au plan P

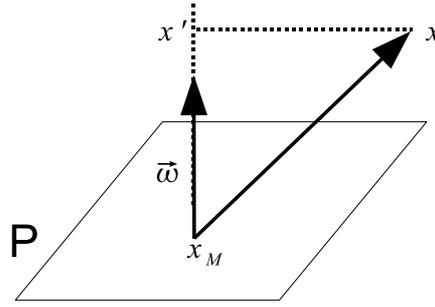
La distance d'un point x au plan P est calculée en projetant x sur la droite (x_M, ω) avec le vecteur ω . On a alors :

$$(x - x_m) \cdot \frac{\omega}{\|\omega\|} = x' - x_M = d(x, P) \quad (8.15)$$

Donc en introduisant b dans l'équation (8.15), on obtient :

$$d(x, P) = \frac{x \cdot \omega + b}{\|\omega\|} \quad (8.16)$$

C'est une distance signée : le signe indique que x est du même côté que ω (> 0) ou de l'autre côté (< 0).

FIG. 8.4 – Distance d'un point x par rapport à un plan P .

c) Le plan séparateur

Définition 8.2 Soit (x_i, y_i) une suite de points et leur classe (-1 ou $+1$) respective. On dit que P est un plan séparateur si :

$$y_i(x_i \cdot \omega + b) \geq 0 \quad (8.17)$$

où ω et b caractérisent le plan P .

Remarque 8.1 La suite de points étant finie, s'il existe P tel que l'équation (8.17) soit vérifiée alors l'inégalité est stricte puisqu'il existe $\varepsilon > 0$ tel que $y_i(x_i \cdot \omega + b) \geq \varepsilon$.

En divisant l'équation (8.17) par ε , on trouve alors que :

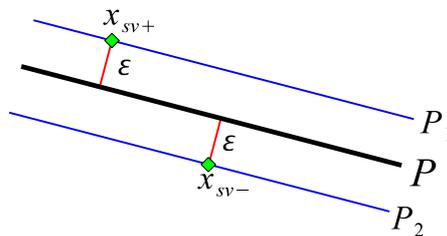
$$y_i(x \cdot \omega' + b') \geq 1 \quad (8.18)$$

avec ω' et b' caractérisant P .

d) Vecteurs supports et calcul de la marge

Comme le montre la figure 8.5, si les données sont linéairement séparables, il existe deux plans parallèles P_1 et P_2 , chacun d'eux intersectant au moins un point noté x_{sv+} ou x_{sv-} selon sa disposition par rapport à P . Ces points x_{sv+} et x_{sv-} sont appelés les **Vecteurs Supports (VS)**.

Le plan P est le plan médian, la distance entre les deux plans P_1 et P_2 est donc 2ε et la marge, mesurée avec la fonction $f'(x) = x \cdot \omega' + b'$, est égale à $\frac{2}{\|\omega\|}$.

FIG. 8.5 – Illustration du plan médian P , des deux plans parallèles, de la marge ε ainsi que des vecteurs supports.

Définition 8.3 Un plan P est dit optimal s'il maximise la marge.

Théorème 8.1 Si la base d'apprentissage \mathcal{A}_n contient n éléments séparés par un plan à marge maximale et si SV est le nombre de vecteurs supports définissant ce plan, l'espérance de la probabilité d'erreur sur l'ensemble d'apprentissage $\mathbb{P}_e(n)$ est bornée par :

$$\mathbb{E}(\mathbb{P}_e(n)) \leq \frac{\mathbb{E}(SV)}{n} \quad (8.19)$$

8.7.3 Marge optimale dans le cas séparable

La recherche d'un hyperplan séparateur à marge maximale revient à résoudre le problème d'optimisation sous contraintes :

$$\begin{cases} \operatorname{argmax} \frac{1}{2} \|\omega\|^2 \\ \omega, b | y_i(x_i \cdot \omega + b) - 1 \geq 0 \end{cases} \quad (8.20)$$

On calcule le maximum par la méthode des multiplicateurs de Lagrange. Dans cette méthode, les contraintes s'écrivent $\psi(x) = 0$ (sans inégalité). Pour se ramener à ce cas, il suffit de considérer la contrainte sur les vecteurs supports, c'est-à-dire les points x_i tel que $y_i(x_i \cdot \omega + b) = 1$. On obtient alors le nouveau problème d'optimisation sous contraintes :

$$\begin{cases} \operatorname{argmax} \frac{1}{2} \|\omega\|^2 \\ \omega, b | y_i(x_i \cdot \omega + b) = 1, \forall i \in SV \end{cases} \quad (8.21)$$

Le problème d'optimisation sous les contraintes (8.21) est donc ramené au problème d'optimisation suivant :

$$\operatorname{argmin} L(\omega, b, \lambda_j) = \frac{1}{2} \|\omega\|^2 - \sum_{i \in SV} \lambda_i (y_i(x_i \cdot \omega + b) - 1) \quad (8.22)$$

La solution à ce problème est obtenue en minimisant par rapport à ω et b et en maximisant par rapports aux λ_i où les coefficients sont $\lambda_i \geq 0$ et sont appelés les multiplicateurs de Lagrange. Pour cela, on calcule la différentielle de L :

$$\frac{\partial L}{\partial \omega} = \omega - \sum_{i \in SV} \lambda_i y_i x_i = 0 \quad (8.23)$$

$$\frac{\partial L}{\partial b} = \sum_{i \in SV} \lambda_i y_i = 0 \quad (8.24)$$

$$\frac{\partial L}{\partial \lambda_i} = y_i(x_i \cdot \omega + b) - 1 = 0 \quad (8.25)$$

Il faut donc trouver (ω, b, λ_i) tels que $w = \sum \lambda_i y_i x_i$, $\sum \lambda_i y_i = 0$ et $y_i(x_i \cdot \omega + b) = 1, \forall i \in SV$. En réinjectant les termes (8.23) et (8.24) dans l'équation (8.22), on obtient :

$$W(\lambda_i) = \sum_{i \in SV} \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j x_i \cdot x_j \quad (8.26)$$

Nous remarquons que W ne dépend plus que des λ_i , que ω se déduit des λ_i par l'équation (8.23) et que b se déduit par l'équation (8.25).

Théorème 8.2 *Le problème de minimisation 8.22 se ramène au problème de maximisation suivant :*

$$\begin{cases} \operatorname{argmax} W(\lambda_i) = \sum_i \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j x_i \cdot x_j \\ \sum_{i \in SV} \lambda_i y_i x_i = 0, \lambda_i \geq 0, \forall i \in SV \end{cases} \quad (8.27)$$

Le lecteur est renvoyé à [Burges, 1998, Fletcher, 1987] pour la preuve de ce théorème.

L'équation (8.26) peut s'écrire sous la forme vectorielle suivante :

$$W(\Lambda) = \Lambda^T \mathbb{1} - \frac{1}{2} \Lambda^T D \Lambda \quad (8.28)$$

avec $\Lambda = (\lambda_1, \dots, \lambda_n)^T$, $\mathbb{1} = (1, \dots, 1)^T$ et $D_{i,j} = (y_i y_j x_i^T x_j)$ est l'élément $1 \leq (i, j) \leq |SV|$ de la matrice D . On a $\frac{\partial D}{\partial \Lambda} = \mathbb{1} - D \Lambda = 0$ donc $\Lambda = D^{-1} \mathbb{1}$. On peut remarquer que D est une matrice symétrique définie positive et qui est donc inversible.

8.7.4 Erreur de classement minimal : cas linéaire non séparable

La formulation (8.26) du problème d'optimisation possède deux inconvénients :

- lorsque les données d'apprentissage ne sont pas géométriquement séparables, c'est-à-dire lorsqu'il n'existe pas de droite séparant les données, le problème n'a pas de solution,
- elle est très sensible aux vecteurs supports. Comme le montre la figure 8.6(b), le plan P_1 est construit en prenant en compte le vecteur support, nommé *outlier*, alors que le plan P_2 serait un plan séparateur plus adéquat si on assouplissait les contraintes.

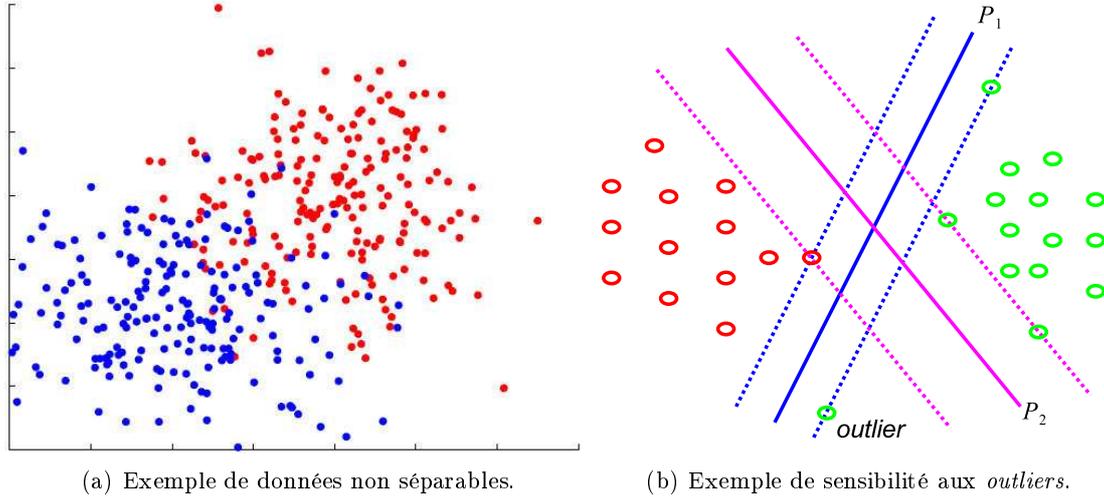


FIG. 8.6 – Sensibilité aux vecteurs supports.

Dans la plupart des problèmes rencontrés en pratique, les données ne sont pas linéairement séparables (voir par exemple la figure 8.6(a)). Il convient de modifier l'approche précédente afin de prendre en compte la possibilité d'observations mal classées. Une solution consiste à reformuler l'équation (8.18) en incorporant des termes d'erreurs $\xi_i \geq 0$ pour mesurer le non respect des contraintes et assouplir ces dernières :

$$y_i(x_i \cdot w + b) \geq +1 - \xi_i, \quad \forall i \in \{1, \dots, n\} \quad (8.29)$$

Une erreur est donc commise sur une observation x_i si le ξ_i correspondant est supérieur à 1. La fonction $\sum_i \xi_i$ constitue un coût caractérisant l'erreur commise sur les éléments de l'ensemble d'apprentissage. Le problème d'optimisation doit être reformulé de manière à minimiser l'erreur de classification conjointement avec la maximisation de la marge. Ceci revient à résoudre le problème de maximisation suivant :

$$\begin{cases} \operatorname{argmax} \frac{1}{2} \|w\|^2 + C \sum_i \xi_i \\ y_i(x_i \cdot w + b) \geq +1 - \xi_i, \quad \xi_i \geq 0 \end{cases} \quad (8.30)$$

où C étant un réel positif préalablement fixé qui contrôle la pénalisation. Plus ce paramètre est petit, plus on pénalise les erreurs.

Définition 8.4 *Un plan dont les paramètres minimisent (8.30) sous les contraintes (8.29) est appelé hyperplan optimal généralisé.*

En effectuant le même raisonnement que pour le cas séparable (section 8.7.3), nous exprimons le lagrangien :

$$L(\omega, b, \lambda_j) = \frac{1}{2} \|\omega\|^2 + C \sum_{i \in SV} \xi_i - \sum_{i \in SV} \lambda_i (y_i(x_i \cdot \omega + b) - 1 + \xi_i) - \sum_{i \in SV} \beta_i \xi_i \quad (8.31)$$

où $\Lambda = (\lambda_1, \dots, \lambda_n)$ et $\beta = (\beta_1, \dots, \beta_n)$ sont les multiplicateurs de Lagrange ($\lambda \geq 0$ et $\beta_i \geq 0 \forall i$).

Les conditions d'annulation des dérivées partielles du lagrangien mènent directement aux relations

vérifiées par l'hyperplan optimal :

$$\sum_{i \in SV} \lambda_i y_i x_i \omega \quad (8.32)$$

$$\sum_{i \in SV} \lambda_i y_i = 0 \quad (8.33)$$

$$\lambda_i + \beta_i = C \quad (8.34)$$

Pour trouver la solution à ce problème, il faut maximiser l'expression suivante :

$$\begin{cases} \text{Argmax } W(\lambda_i) = \sum_i \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j x_i \cdot x_j - \frac{\delta^2}{C} \\ \sum_{i \in SV} \lambda_i y_i x_i = 0 \\ 0 \leq \lambda_i \leq C, \forall i \in SV \end{cases} \quad (8.35)$$

Ce qui est équivalent à la forme vectorielle suivante :

$$W(\Lambda, \delta) = \Lambda^T 1 - \frac{1}{2} \left[\Lambda^T D \Lambda + 2 \frac{\delta^2}{C} \right] \quad (8.36)$$

où $\delta = \max(\lambda_1, \dots, \lambda_n)$ avec les contraintes $\Lambda \geq 0$ et $\Lambda^T Y = 0$.

8.7.5 Séparation hautement non linéaire

Dans le cas où le problème de classification ne peut se résoudre grâce à la recherche d'un hyperplan (avec ou sans prise en compte de l'erreur), nous nous trouvons dans le cas de données hautement non linéaires (géométriquement parlant). Il faut alors se ramener à un problème linéaire en transformant les données à l'aide d'une fonction Φ dans un espace de dimension supérieure.

Soit (x_i, y_i) un ensemble d'apprentissage non linéairement séparable avec $x_i \in \mathbb{R}^m$. Il est toujours possible de trouver $z_i \in \mathbb{R}$ et un scalaire ω_z tels que la relation suivante soit vérifiée :

$$y_i(\omega \cdot x_i + \omega_z \cdot z_i + b) = 1 \quad \forall i \quad (8.37)$$

L'idée de Vapnik est de trouver une fonction qui calcule le z_i en fonction des données d'apprentissage x_i .

La figure 8.7 présente un exemple de transformation de données non linéaires en données linéairement séparables par l'ajout d'une troisième composante z dans cet exemple.

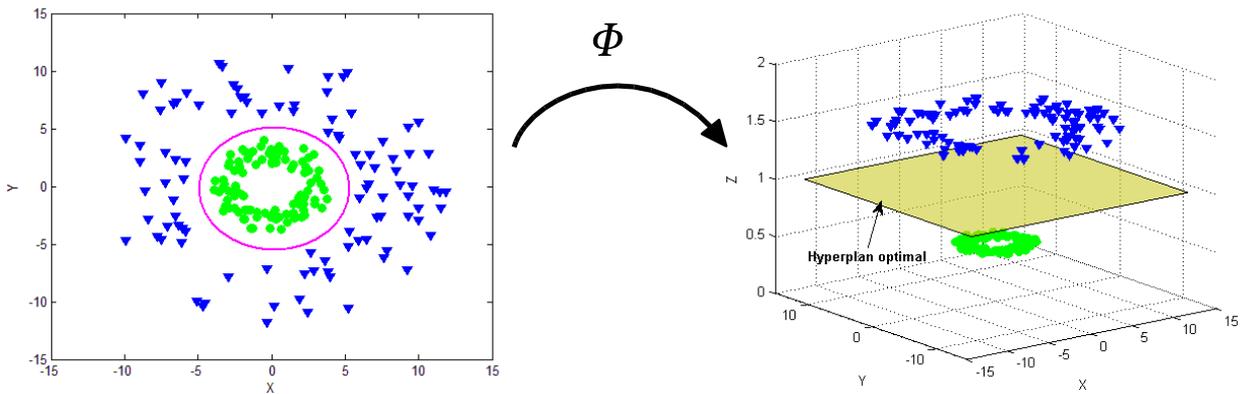


FIG. 8.7 – Exemple de transformation de données non linéaire en données linéairement séparables.

Cette méthode nécessite de connaître la formule de passage vers ce nouvel espace, ce qui n'est pas évident. En réalité, il n'est pas nécessaire de connaître explicitement Φ comme nous allons le voir. [Cortes and Vapnik, 1995] propose l'idée suivante :

- toute fonction $K(u, v)$ symétrique et positive forme un produit scalaire convenable dans l'espace de dimension supérieure obtenu via la fonction Φ . Cette affirmation repose sur la théorie de Hilbert [Courant and Hilbert, 1953] qui affirme que toute fonction symétrique de L^2 peut s'écrire :

$$K(u, v) = K(v, u) = \sum_{i=1}^{\infty} \lambda_i \Phi_i(u) \cdot \Phi_i(v) \quad (8.38)$$

où $\{\Phi_i\}$ est une base de L^2 et λ_i sont tels que :

$$\int K(u, v) \Phi_i(u) du = \lambda_i \Phi_i(v) \quad (8.39)$$

$$\langle K, \Phi_i \rangle_{L^2} = \lambda_i \Phi_i \quad (8.40)$$

λ_i est la valeur obtenue lorsque l'on projette K sur Φ_i , Φ_i étant les vecteurs propres et λ_i les valeurs propres associées.

Afin de garantir que les valeurs propres de l'équation (8.38) soient positives, il est nécessaire et suffisant (Théorème de Mercer) que la condition :

$$\iint K(u, v) g(u) g(v) dudv > 0$$

soit satisfaite pour toute fonction g de L^2 .

- Pour calculer le plan séparateur dans cet espace, il n'est donc pas nécessaire de calculer les $\Phi(x_i)$ mais seulement les produits scalaires $\langle \Phi(x_i), \Phi(x_j) \rangle = K(x_i, x_j)$.

En appliquant l'équation (8.35) dans l'espace engendré par Φ , on obtient :

$$\begin{aligned} W(\lambda_j) &= \sum_i \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j \langle \Phi(x_i), \Phi(x_j) \rangle \\ &= \sum_i \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j K(x_i, x_j) \end{aligned} \quad (8.41)$$

sous les contraintes $\sum_i \lambda_i y_i = 0$ et $0 \leq \lambda_i \leq C$ où K est un noyau tel que $K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$. Et l'équation de l'hyperplan optimal est :

$$\omega \cdot \Phi(x) + b = \sum_{i \in SV} \lambda_i y_i K(x_i, x) + b \quad (8.42)$$

où $b = \frac{1}{2} \sum_{i \in SV} \lambda_i y_i [K(x_i, x_{sv-}) + K(x_i, x_{sv+})]$.

Règle de décision

Pour classifier une nouvelle donnée, il suffira de regarder la position de cette donnée rapport au plan P , c'est-à-dire le signe de la fonction de décision suivante :

$$I(x) = \text{sign} \left(\sum_{i \in SV} \lambda_i y_i K(x_i, x) + b \right) \quad (8.43)$$

Quelques exemples de noyaux

Différents noyaux K ont été étudiés dans la littérature. Dans la pratique, il faut combiner des noyaux simples pour en obtenir de plus complexes adaptés à un problème. De plus, le choix du noyau se fait généralement d'une manière empirique lors de tentative de type essai-erreur. Voici une liste non exhaustive des noyaux les plus courants :

- Linéaire :

$$K(x, x') = \langle x, x' \rangle$$

- Polynomial :
- les noyaux polynomiaux de degré q sont de la forme :

$$K(x_1, x_2) = (c + \langle x_1, x_2 \rangle)^q$$

- Gaussien (*Radial Basis Function* ou RBF) :
- les noyaux gaussiens sont des noyaux de type *radial*, indiquant qu'ils dépendent de la distance $\|x_1 - x_2\|$ entre les observations. Ces noyaux sont définis par :

$$K(x, x') = e^{-\frac{\|x-x'\|^2}{2\sigma^2}}$$

où $2\sigma^2$ est appelé *largeur de bande*.

- Sigmoidal :
- Le noyau sigmoïdal est défini comme :

$$K(x, x') = \tanh(a\langle x, x' \rangle - b)$$

Ce noyau dépend de deux paramètres a et b qu'il faut fixer pour permettre la mise en oeuvre.

Un exemple dans la pratique :

Soit $x = (x_1, x_2) \in \mathbb{R}^2$ et $\Phi(x) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)$.

Dans ce cas, \mathcal{H} est de dimension 3 et le produit scalaire s'écrit :

$$\begin{aligned} \langle \Phi(x), \Phi(x') \rangle &= x_1^2x_1'^2 + 2x_1x_2x_1'x_2' + x_2^2x_2'^2 \\ &= (x_1x_1' + x_2x_2')^2 \\ &= \langle x, x' \rangle^2 \\ &= K(x, x') \end{aligned}$$

8.7.6 Classification multi-classes avec les SVM

Nous avons vu jusqu'ici que l'algorithme des SVM réalise une classification binaire. Mais il peut fonctionner comme un classifieur multi-classes par la combinaison de plusieurs classifieurs binaires pour chaque paire de classes possible. Les probabilités pour qu'un pixel x appartienne à la classe i sont calculées. On peut ainsi définir une fonction discriminante linéaire telle :

$$g_i(x) = w_i^t y \quad \forall i = 1, \dots, M \quad (8.44)$$

où w est un vecteur correcteur qui permet d'ajuster les fonctions discriminante en augmentant les valeurs pour la bonne classe et en la décrémentant pour les autres. Prenons le cas où la j^{me} fonction discriminante $g_j(x)$ présente une grande erreur alors que le pixel appartient à la classe i . On définit w de la façon suivante :

$$\begin{aligned} w_i' &= w_i + c.y \\ w_j' &= w_j - c.y \end{aligned} \quad (8.45)$$

où c joue de valeur de correction. Cette procédure d'ajustement est réalisée autant de fois que nécessaire pour chacun des pixels de l'ensemble d'apprentissage pour obtenir la solution. La décision est prise en prenant la plus grande valeur des g_i . C'est la bibliothèque LIBSVM [Chang and Lin, 2001] qui sera utilisée pour réaliser les classifications par la méthode des SVM et mis en oeuvre en utilisant un noyau de fonctions à base radiales.

Choix du classifieur et des données d'entrée

Nous comparons les performances de classifieurs vus au chapitre 8 pour un problème type de classification des sols. Nous décrivons la banque des classes utilisées et nous montrons comment coder l'information d'un pixel afin d'obtenir une classification correcte en combinant l'information spectrale et l'information de texture. Par ailleurs, l'utilisation du classifieur le plus performant permet d'obtenir directement le masque de la végétation haute.

9.1 Mesure de la séparabilité des classes

Nous appliquons maintenant les algorithmes présentés au chapitre 8 aux données Ikonos à notre disposition. Évidemment, nous ne disposons pas de la bande proche infrarouge (PIR) indispensable pour les calculs des indices de végétation. Comme dans la seconde partie de cette thèse, nous devons trouver une stratégie alternative pour caractériser la signature de la végétation dans les images Ikonos. Pour mesurer l'influence des informations de texture et des bandes spectrales en terme de qualité de la classification, nous utilisons la distance de Jeffries-Matusita (JM). Cette mesure donne la distance moyenne entre deux classes [Wacker, 1971] et indique la séparabilité de ces classes. Pour une distance entre deux classes i et j , elle est définie par :

$$JM_{ij} = 2(1 - e^{B_{ij}}) \quad (9.1)$$

où B est la distance de Bhattacharyya [Kailath, 1967] définie par :

$$B_{ij} = \frac{1}{8}(m_i - m_j)^t \left(\frac{\Sigma_i + \Sigma_j}{2} \right)^{-1} (m_i - m_j) + \frac{1}{2} \ln \left(\frac{|(\Sigma_i + \Sigma_j)/2|}{\sqrt{|\Sigma_i|}\sqrt{|\Sigma_j|}} \right) \quad (9.2)$$

m_i , m_j sont les vecteurs des moyennes des classes i et j , et Σ_i , Σ_j les matrices de covariances des classes i et j . La distance JM prend des valeurs comprises dans l'intervalle $[0; 2]$, et renseigne sur la séparabilité, en un sens statistique, entre deux classes. Une valeur au dessus de 1,9 indique une très bonne séparabilité. Une valeur inférieure à 1 indique une faible séparabilité. Dans ce cas, il faut envisager la fusion de certaines classes afin d'augmenter la séparabilité des nouvelles classes.

Il faut donc sélectionner les données d'apprentissage de façon à obtenir la plus grande distance JM en moyenne. Cette distance moyenne est donnée par :

$$d_{average} = \sum_{i=1}^M \sum_{j=1}^M p(\omega_i)p(\omega_j)JM_{ij} \quad (9.3)$$

pour M classes. Les $p(\omega_i)$ sont les probabilités *a priori* pour chaque classe. Dans la suite de cette thèse, nous désignerons sous le terme de séparabilité des classes la mesure de la distance JM entre ces classes.

9.2 Choix des ensembles d'apprentissage

Une bonne connaissance du terrain est utile pour choisir correctement les ensembles d'apprentissage et définir des classes pertinentes. Une étude approfondie de l'image et des couvertures des sols des atolls des Tuamotu (généralement, nous remarquons le même type de couverture de sol) nous conduit à définir un jeu de neuf classes. Le tableau 9.1 décrit les classes utilisées que nous avons utilisées pour la classification des images Ikonos. La figure 9.1 montre les différentes classes superposées à l'image originale de l'atoll. La classe végétation (**Vg**) regroupe toute la végétation composée très majoritairement de cocotiers.

DW	Eaux Profonde	Deep Water
MW	Eaux Profondeur intermédiaire	Med deep Water
SW	Eaux peu profondes	Shallow Water
Vg	Végétation	Vegetation
WS	Sable blanc	White Sand
DC	Corail Sec	Dry Coral
BC	Corail sombre	Black Coral
Wa	Vagues	Waves
AS	Piste d'atterrissage	AirStrip

TAB. 9.1 – Acronyme des différentes classes utilisées.

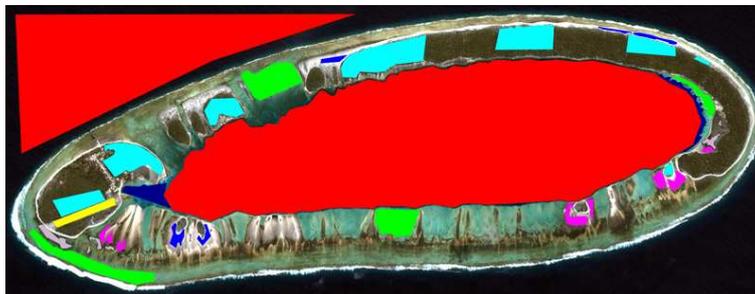


FIG. 9.1 – Localisation des différentes classes sur l'image originale.

La figure 9.2 présente les différents histogrammes calculés sur les régions qui définissent les classes citées dans le tableau 9.1.

Pour chaque classe du tableau 9.1, un nombre de petites régions caractéristiques sont choisies empiriquement. Afin d'obtenir des ensembles d'apprentissage de même taille, un nombre égal de pixels y sont tirés aléatoirement. À partir de cet ensemble de données d'apprentissage, il est possible d'extraire diverses statistiques telles que la séparabilité des classes, histogrammes de chaque classes, etc. La figure 9.3 montre les histogrammes des ensembles d'apprentissages. Le tableau 9.2 présente la séparabilité des classes calculée sur les informations spectrales. Nous remarquons que la séparabilité moyenne totale des classes est de 1,816 si seules les informations spectrales sont prises en compte. En effet, le corail sec (DC) et la piste (AS) ont une séparabilité totale moyenne très proche ce qui engendre une erreur de classification des pixels pour ces deux classes. Seules deux classes, eaux profondes (DW) et eaux peu profondes (SW), présentent une séparabilité moyenne totale supérieure à 1,9 alors que toutes les autres classes ont des valeurs en dessous de 1,86. Ceci signifie que seuls les pixels des classes DW et SW sont correctement étiquetés. Nous verrons par la suite comment compléter l'information pour augmenter la séparabilité et obtenir une classification correcte.

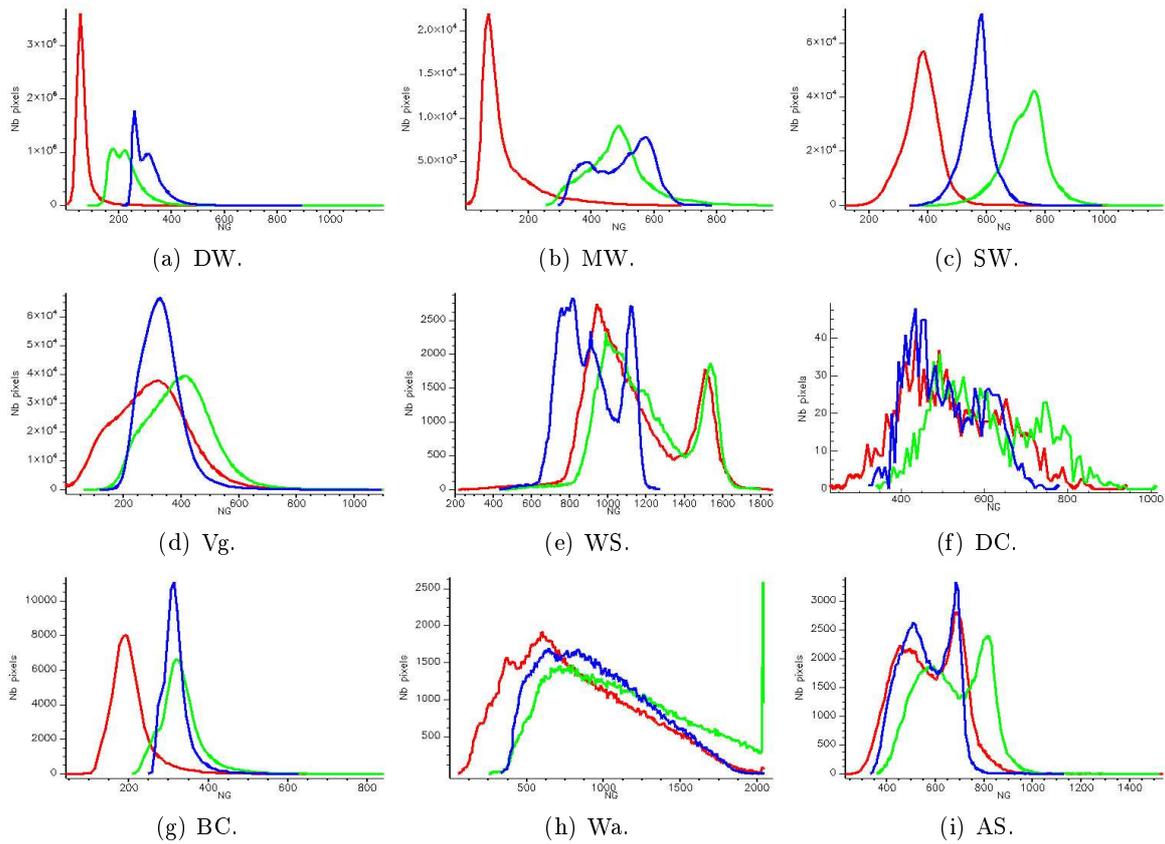


FIG. 9.2 – Histogrammes des neuf classes.

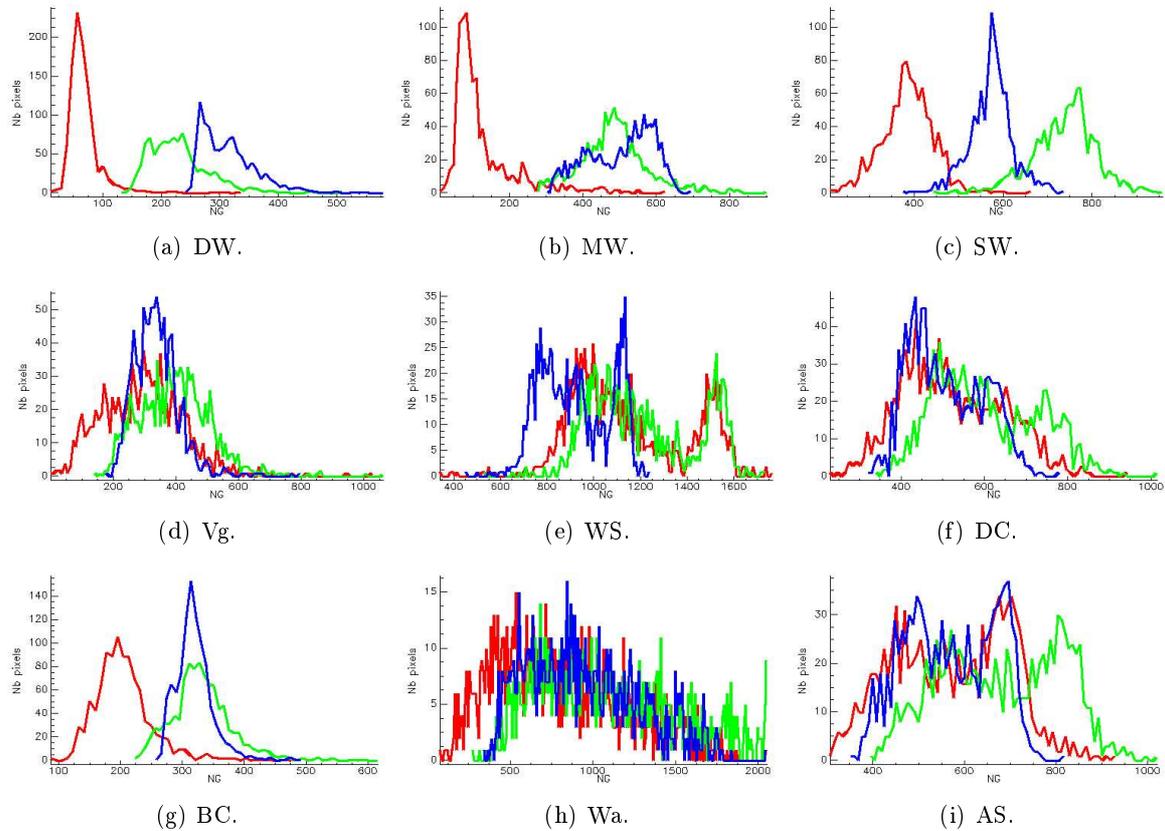


FIG. 9.3 – Histogrammes des ensembles d'apprentissage des neuf classes.

	DP	MW	SW	Vg	WS	DC	BC	Wa	AS	
DP	0									
MW	1,8453926	0								
SW	2,0000000	1,8298295	0							
Vg	1,9822554	1,8568069	1,9983129	0						
WS	1,9969410	1,9966158	1,9999985	1,9774774	0					
DC	1,9990931	1,9867614	1,9999803	1,5855699	1,6257159	0				
BC	1,9770562	1,8307417	1,9983443	1,2267569	1,9900571	1,8836002	0			
Wa	1,9402163	1,5159617	1,9703415	1,7617485	1,7346755	1,8098918	1,8627966	0		
AS	1,9990474	1,9893184	1,9999865	1,7437476	1,4638695	0,3491250	1,9125185	1,7395199	0	Moy. Totale
Moy.	1,9675002	1,8564285	1,9745992	1,7665844	1,8481688	1,6549672	1,8352339	1,7918940	1,6496416	1,8161131

TAB. 9.2 – Séparabilité des classes calculée sur les bandes spectrales RGB.

9.3 Influence des informations de texture sur la séparabilité des classes

Analyse en Composante Principale des bandes spectrales

Nous devons dans un premier temps préparer les données avant d'effectuer le calcul des informations de texture. Pour cela, nous effectuons une ACP l'image RGB comme décrit dans la sous-section 5.2.3. La figure 9.4 montre un exemple de calcul de l'ACP sur une image RGB.

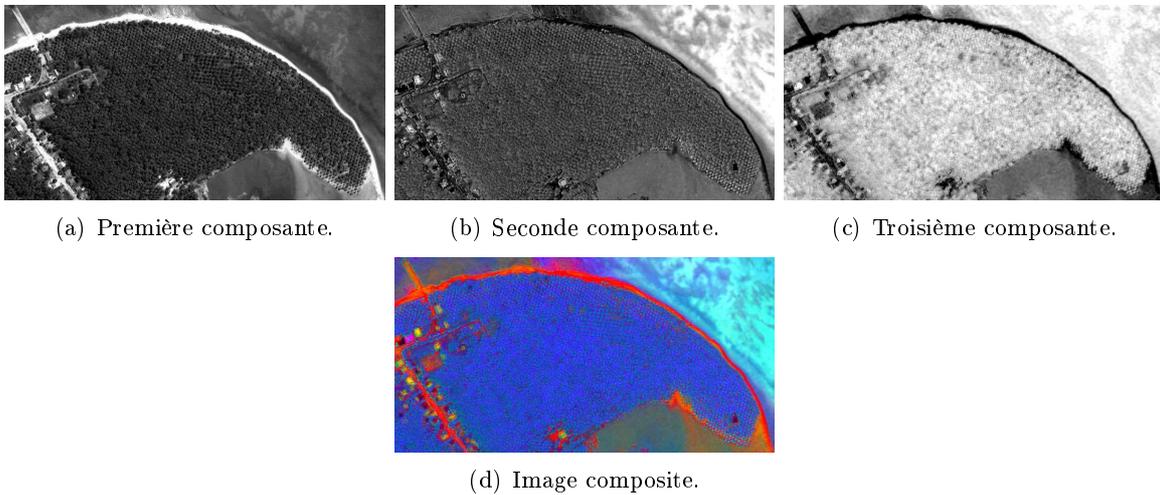


FIG. 9.4 – Résultat de l'ACP sur l'image RGB originale. L'image 9.4(d) est une image synthétique composée de la première composante pour le rouge, la seconde composante pour le vert et la dernière pour bleu.

La figure 9.5 présente l'histogramme des trois composantes de l'ACP sur l'image RGB. Le tableau 9.3 résume les statistiques telles que valeurs minimales, valeurs maximales, moyennes, écart types et valeurs propres pour chaque composante. La première composante possède quasiment toute l'information (inertie de 96%) et que ces composantes principales sont décorrélées entre elles. L'analyse en composante principale est donc un outil efficace pour fusionner les informations contenues dans les trois bandes spectrales en un seul canal. C'est donc la première composante qui sera utilisée pour calculer les informations de texture.

Calcul des descripteurs de texture

Comme l'a montré le tableau 9.2, les informations ne sont pas suffisantes pour discriminer correctement les classes. Pour cela, nous allons calculer des informations de texture. Les descripteurs de texture utilisés [Haralick et al., 1973, Haralick, 1979, Anys et al., 1994, Puissant et al., 2005] sont détaillés en annexe H. Le calcul de ces descripteurs se fait à partir d'une matrice de cooccurrence. Cette matrice permet d'estimer le nombre d'occurrence d'un couple de niveaux de gris (x_u, x_v) séparé d'une distance d et situé sur une droite faisant un angle θ par rapport à l'axe horizontal. Le choix de la fenêtre de calcul dépend des structures présentes dans les images. Comme nous voulons classer les différents des

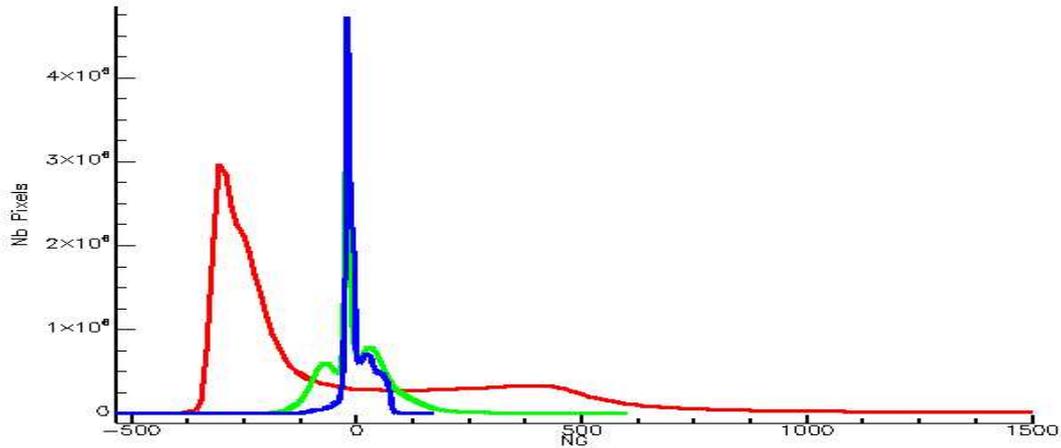


FIG. 9.5 – Histogramme des trois composantes après l'ACP.

	Min	Max	Mean	StDev		ACP1	ACP2	ACP3
ACP 1	-463,4273	2912,4709	2,7333	396,5679	ACP 1	1		
ACP 2	-409,4496	599,1708	0,1824	69,3021	ACP 2	-0,0018	1	
ACP 3	-534,0936	167,5140	0,1900	33,3275	ACP 3	-0,0039	-0,0016	1

(a) Bases.

(b) Correlation.

Bandes	Valeurs Propres	Vecteurs Propres		
ACP 1	157266,1463	1	-0,0003	-0,0003
ACP 2	4802,7653	-0,0003	-1	0,0010
ACP 3	1110,7009	-0,0003	-0,0010	-1

(c) Valeurs et vecteurs propres.

TAB. 9.3 – Les statistiques des composantes principales.

types de plantations visibles sur les images Ikonos, il est nécessaire de connaître l'organisation spatiale de différents types de plantations de cocotiers. Les descripteurs de texture utilisés sont : la variance, l'homogénéité, le contraste, l'entropie, le second moment angulaire, la corrélation. En plus de ces six descripteurs, nous prenons en compte la moyenne (eq. (H.1)) ainsi qu'une autre mesure : la dissimilarité définie à partir de la matrice de cooccurrence :

$$\text{Dissimilarité} = \sum_{i=0}^{NG-1} \sum_{j=0}^{NG-1} P(i, j) |i - j| \quad (9.4)$$

Les plantations pouvant avoir une orientation quelconque, il ne faut pas privilégier une orientation particulière. Les descripteurs de texture sont calculés pour quatre orientations : 0° , 45° , 90° et 135° de manière symétrique et on calcule la moyenne de ces quatre orientations. Par la suite, ce sont les descripteurs moyens de texture qui sont utilisés d'abord pour le calcul de la séparabilité des classes puis pour la classification.

Pour fixer la taille de la fenêtre glissante, nous avons utilisé les connaissances *a priori* des plantations données par le Service du Développement Rural (SDR) de la Polynésie Française. Elles ont permis de mettre en évidence cinq types de plantations artificielles (pour lesquels une organisation spatiale est visible) dans les cocoteraies (voir la figure 9.6). Dans le cas des plantations en carré ou quinconce, on a $A = B = 8$ m soit avec une résolution du pixel au sol de 80 cm, on obtient une distance de 10 pixels. Dans le cas d'une organisation en quinconce équilatérale, on a $B = \frac{\sqrt{3}}{2} \times A$. Donc avec $A = 8$ m on obtient $B = 6,93$ m = 8,66 pixels soit 9 pixels. Et enfin, dans le cas d'une organisation spatiale en quinconce équilatéral avec couverture totale, on a $A = \frac{\sqrt{3}}{2} \times D \iff D = \frac{1}{0,87} \times A$, et $B = \frac{\sqrt{3}}{2} \times A$. Pour $A = 8$ m, on obtient $B = 9$ pixels et $D = 9,24$ m = 11,55 pixels soit 12 pixels. Grâce à ces informations sur l'organisation spatiale des cocoteraies, la taille de la fenêtre utilisée dans le calcul de la matrice de cooccurrence est choisi à 17×17 pixels pour capturer des motifs représentatifs des

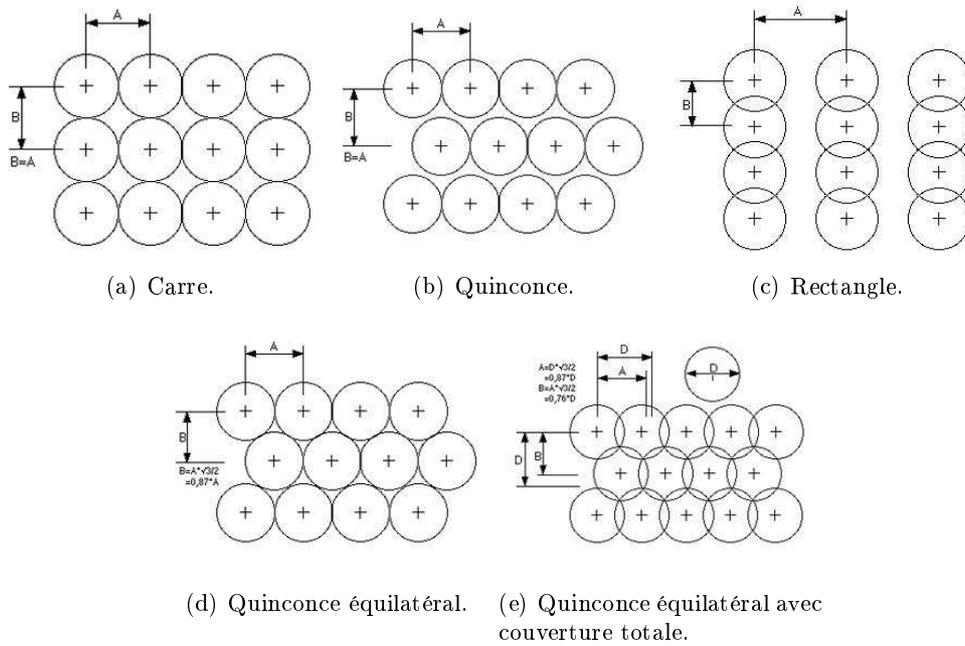
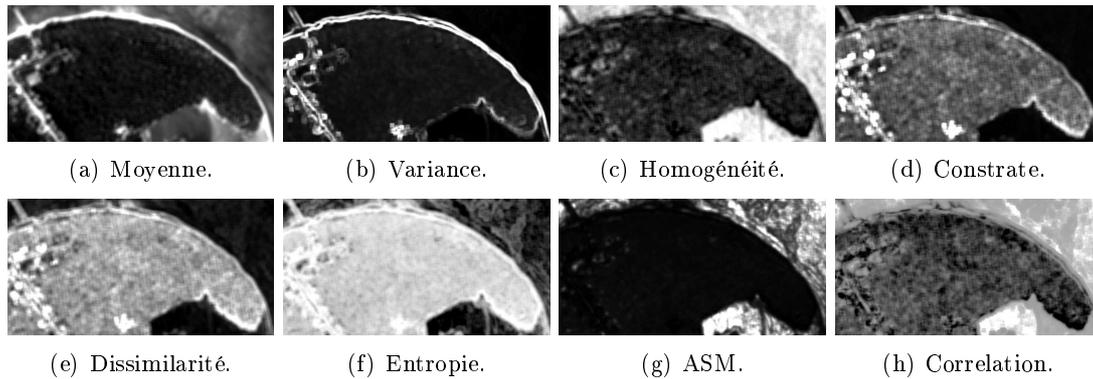


FIG. 9.6 – Différentes organisations spatiales des cocoteraies artificielles.

différentes configurations. La figure 9.7 présente les différents descripteurs de texture calculés avec une glissante de 17×17 pixels.

FIG. 9.7 – Les moyennes descripteurs de texture calculés pour une fenêtre de 17×17 pixels.

Analyse de la séparabilité des classes

Les tableaux 9.2, 9.4 et 9.5 présentent la séparabilité des classes en prenant en compte uniquement l'information spectrale (tableau 9.2), l'information texturale (tableau 9.4) et enfin les informations texturale et spectrale (tableau 9.5). Nous avons vu dans la sous-section 9.2 que l'information spectrale à elle seule ne permettait pas d'obtenir une classification satisfaisante. L'apport de l'information de texture permet maintenant de classifier correctement les classes. En effet, les classes MW, Vg, WS, DC, BC, Wa et AS ont une séparabilité supérieure à 1,9. On peut néanmoins remarquer une légère baisse de la qualité pour la classe SW mais celle-ci reste marginale car elle est toujours supérieure à 1,9. Finalement, on passe d'une séparabilité moyenne totale de 1,816 avec l'information spectrale à 1,972 avec l'information de texture sachant que toutes les classes sont correctement détectées. Pour finir, on peut combiner les deux informations (spectrale et texturale) et monter la séparabilité à 1,996 c'est-à-dire très proche de la valeur idéale de 2 (pas d'erreur de classification).

Le tableau 9.6 résume les différentes séparabilités moyennes totales calculées en prenant en compte uniquement l'information spectrale, l'information de texture et la combinaison des deux. Ces résul-

	DP	MW	SW	V _g	WS	DC	BC	Wa	AS	
DP	0									
MW	1,9913981	0								
SW	1,9999998	1,9856455	0							
V _g	2,0000000	1,9999995	1,9999994	0						
WS	1,9999972	1,9997802	1,9933240	1,9998315	0					
DC	1,9985108	1,8004663	1,6526227	1,9999954	1,9871047	0				
BC	1,9719448	1,9175206	1,9998444	1,9999705	1,9994084	1,9839007	0			
Wa	2,0000000	1,9999691	1,9999330	1,9976147	1,9551461	1,9999034	1,9997111	0		
AS	1,9999994	1,9975136	1,9677966	1,9977006	1,8948737	1,9309355	1,9983540	1,9948442	0	Moy,Totale
Moy.	1,9952313	1,9615366	1,9498957	1,9993889	1,9786832	1,9191799	1,9838318	1,9933902	1,9727522	1,9726544

TAB. 9.4 – Séparabilité des classes calculée sur les textures.

	DP	MW	SW	V _g	WS	DC	BC	Wa	AS	
MW	1,9970800	0								
SW	2,0000000	1,9985055	0							
V _g	2,0000000	2,0000000	2,0000000	0						
WS	2,0000000	1,9999977	2,0000000	1,9999836	0					
DC	1,9999988	1,9990003	1,9999990	1,9999984	1,9914444	0				
BC	1,9999713	1,9930349	1,9999766	1,9999979	1,9998843	1,9971027	0			
Wa	2,0000000	1,9999911	1,9999959	1,9995257	1,9927203	1,9999892	1,9999487	0		
AS	2,0000000	1,9999837	1,9999999	1,9990011	1,9391216	1,9646150	1,9999365	1,9991644	0	Moy,Totale
Moy.	1,9996288	1,9984466	1,9998096	1,9998133	1,9903940	1,9940185	1,9987316	1,9989169	1,9877278	1,9963875

TAB. 9.5 – Séparabilité des classes calculée sur les textures et les bandes spectrales.

tats montrent donc l'intérêt d'intégrer les informations texturales aux informations spectrales pour la discrimination entre classe pour les images Ikonos de nos atolls.

Tab.	DP	MW	SW	V _g	WS	DC	BC	Wa	AS	Moy>Total
9.2	1,9675002	1,8564285	1,9745992	1,7665844	1,8481688	1,6549672	1,8352339	1,7918940	1,6496416	1,8161131
9.4	1,9952313	1,9615366	1,9498957	1,9993889	1,9786832	1,9191799	1,9838318	1,9933902	1,9727522	1,9726544
9.5	1,9996288	1,9984466	1,9998096	1,9998133	1,9903940	1,9940185	1,9987316	1,9989169	1,9877278	1,9963875

TAB. 9.6 – Récapitulatif des séparabilité des classes.

Comparaison des performances des algorithmes de classification

Nous avons mesuré les performances de quatre classifieurs (voir le chapitre 8) : distance minimale, distance de Mahalanobis, bayésien et SVM avec un noyau de fonctions à base radiales de variance γ . Les performances d'un classifieur sont déterminées empiriquement, en sélectionnant (de préférence aléatoirement) un échantillon de pixels de la classification et en vérifiant leurs étiquettes par rapport aux classes des ensembles d'apprentissage. Pour chaque classe, le taux des pixels correctement classifiés peut être estimé ainsi que les pixels mal classifiés. Ces résultats se présentent sous la forme de matrices de confusions (ou d'erreurs) [Jensen, 1986]. Les valeurs de la diagonale de la matrice représentent le nombre de pixels correctement classifiés. Les autres éléments (i, j) de la matrice indiquent le nombre de pixels finalement étiquetés avec la classe i alors qu'ils sont supposés appartenir à la classe j . Une matrice de confusion est idéalement diagonale lorsqu'il n'y a pas d'erreur de classification. Le coefficient κ permet alors de mesurer l'efficacité d'un classifieur [Hudson and Ramm, 1987, Congalton and Green, 1999] :

$$\kappa = \frac{N \sum_k x_{kk} - \sum_k x_{\Sigma k} x_{k\Sigma}}{N^2 - \sum_k x_{\Sigma k} x_{k\Sigma}} \quad (9.5)$$

où N est le nombre de pixel, les coefficients x_{kk} sont les termes diagonaux de la matrice de confusion, $x_{\Sigma k}$ est la somme des éléments d'une colonne et $x_{k\Sigma}$ est la somme des éléments d'une ligne. Plus κ est proche de 1, meilleure est la précision de la classification.

Le tableau 9.7 montre les performances globales des classifieurs pour chaque classe ainsi que le coefficient κ . Nous observons que la classification par distance minimale n'est pas adaptée à notre étude et fournit seulement une performance de classification de 63,7% avec une valeur pour κ faible de l'ordre de 0,5. La classification par distance de Mahalanobis ainsi que la classification bayésienne donne des résultats corrects avec une performance respective de 84,27% et 91,92%. La classification

	Distance Minimale		Mahalanobis		ML		SVM	
	Performance		Performance		Performance		Performance	
	Product	User	Product	User	Product	User	Product	User
DW	93,90%	79,37%	91,10%	90,56%	93,00%	97,69%	97,50%	98,39%
MW	68,50%	88,96%	73,20%	88,30%	94,90%	82,38%	95,60%	95,50%
SW	95,30%	84,11%	98,10%	98,20%	97,80%	98,19%	98,30%	98,01%
Vg	49,50%	49,80%	96,60%	94,52%	97,50%	99,69%	98,30%	99,29%
WS	67,10%	65,72%	83,90%	97,22%	86,80%	96,77%	98,20%	98,69%
DC	40,70%	41,45%	73,90%	70,92%	92,10%	78,79%	91,00%	88,61%
BC	77,90%	64,27%	79,30%	70,36%	83,10%	99,88%	98,80%	98,41%
Wa	28,30%	45,28%	76,80%	93,43%	97,50%	88,80%	95,40%	97,05%
AS	52,10%	48,24%	85,50%	66,28%	84,60%	91,36%	88,60%	88,07%
Perf. globale	63,7000%		84,2667%		91,9222%		95,7444%	
Coeff. κ	0,5916		0,8230		0,9091		0,9521	

TAB. 9.7 – Comparaison des classifieurs.

par SVM fourni la meilleure classification avec une performance approchant les 96%. La figure 9.8 montre le résultat de la classification des sols de l'atoll de Vahitahi par SVM.

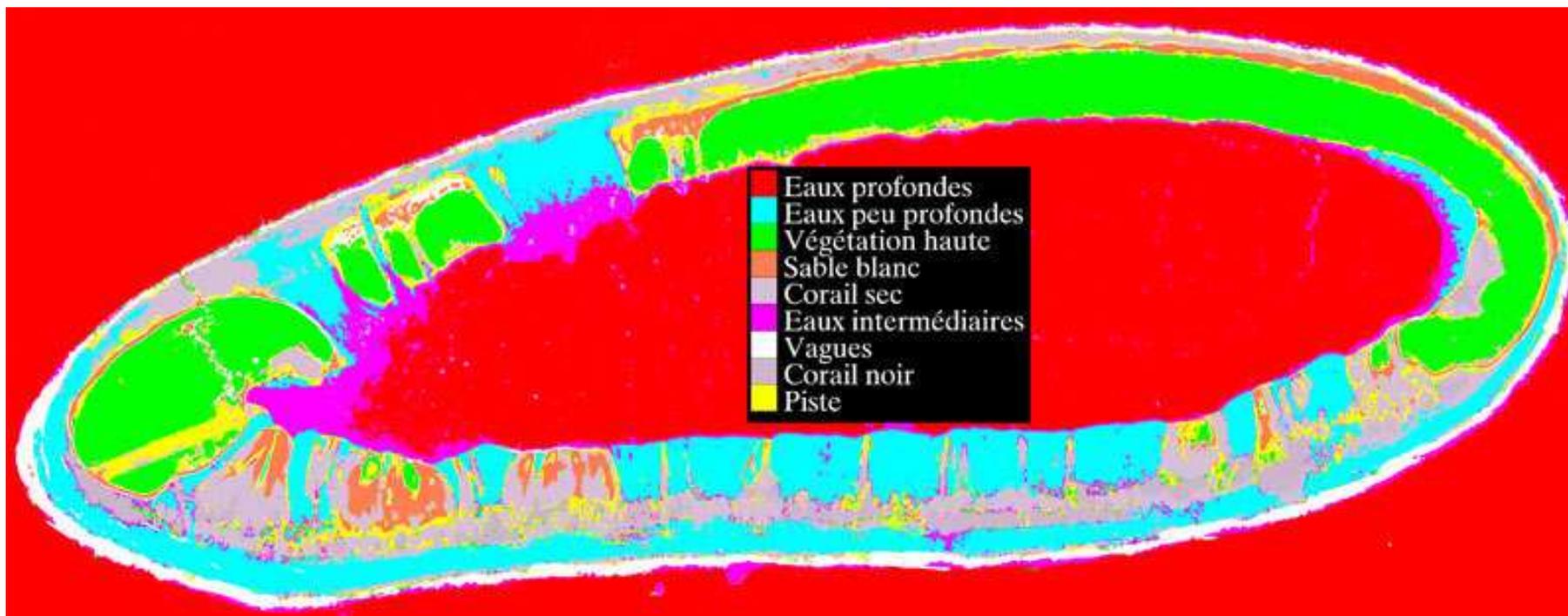


FIG. 9.8 – Classification par SVM de l'occupation des sols de l'atoll de Vahitahi.

Conclusion

Nous avons vu l'influence des informations texturales dans la qualité de la classification et qu'il est nécessaire de les combiner aux informations spectrales pour obtenir de bonnes classifications des sols. Ce qui se distingue dans ce chapitre, par rapport à ce qui est fait lors de la création des masques de végétation la sous-section 5.2.2, est la distinction directe de la végétation haute et de la végétation basse. Il n'est donc plus nécessaire de réaliser un seuillage dont la valeur de seuil est choisi par l'analyse d'un profil. Nous avons montré que la classification par SVM a les meilleurs performances. Nous avons donc utilisé ce classifieur pour produire des cartes thématiques de l'occupation des sols des atolls. Ces cartes sont exploitées actuellement par le SDR pour estimer les surfaces de cocoteraies. L'annexe I présente les résultats de l'estimation des surfaces de cocoteraies supérieures à 1 hectare sur les atolls de Vahitahi, Vairaatea, Nukutavake et Fangatau, le tableau 9.8 résume les surfaces estimées pour chacun de ces atolls.

Nom de l'atoll	Surface en Hectares
Fangatau	480,977
Nukutavake	331,576
Vahitahi	257,972
Vairaatea	243,911

TAB. 9.8 – Surfaces de cocoteraies estimées.

Caractérisation des cocoteraies à partir de la texture

Dans le chapitre précédent, nous avons vu une méthode efficace pour classifier les sols et par conséquent identifier les cocoteraies. Nous souhaitons maintenant affiner l'étude de ces zones en caractérisant l'organisation spatiale des cocotiers. En particulier, il s'agit de caractériser les cocoteraies en trois classes :

- sauvage : la cocoteraie dans laquelle aucune organisation spatiale ne se distingue,
- artificielle : la cocoteraie plantée par l'homme avec un espacement régulier et un alignement,
- intermédiaire : la cocoteraie artificielle laissée à l'abandon.

L'intérêt de cette caractérisation des cocoteraies est principalement économique : outre le recensement des cocotiers, elle permet d'évaluer le potentiel d'exploitabilité des cocoteraies en aménageant et structurant l'espace de façon à optimiser la production.

Ce chapitre décrit une première approche en utilisant les valeurs spectrales des pixels comme vecteur d'entrée d'un classifieur. Pour cela, nous devons résoudre deux problèmes :

- quelles sont les informations spectrales nécessaires pour permettre une classification correcte ?
- quel classifieur est le plus adapté pour caractériser les cocoteraie ?

À partir d'un ensemble d'apprentissage, nous déterminons les informations spectrales nécessaires pour obtenir une bonne classification. Plusieurs classifieurs sont étudiés pour déterminer leur efficacité.

Pour illustrer cette étude, nous nous intéressons à un autre secteur de l'atoll de Vahitahi. Ce secteur, situé au Nord-Est (voir figure 3.10), a la propriété d'avoir différents types de plantations disponibles sur une portion relativement faible de l'image.

10.1 Codage de l'information

10.1.1 Ensemble d'apprentissage

L'ensemble d'apprentissage est constitué de régions qui décrivent chaque classe de cocoteraies. Afin d'éviter d'avoir une matrice de covariance singulière [Swain, 1978], il est nécessaire d'avoir $100 \times N$ (où N est le nombre de bandes) pixels d'apprentissage pour chaque classe. Les données d'apprentissage sont choisies par tirage aléatoire de 1500 pixels dans chaque région. La figure 10.1 présente les trois classes de type de plantation listées ci-dessous :

- 10.1(a) : plantation naturelle dans laquelle nous ne distinguons aucune organisation spatiale (Natural),
- 10.1(b) : plantation artificielle à basse densité (Artificial LD) avec une organisation spatiale. L'espacement entre deux arbres est élevé et cette configuration permet de distinguer le sol,
- 10.1(c) : plantation artificielle à densité moyenne (Artificial MD). Dans une même surface que la classe précédente, nous avons une quantité plus grande d'arbres (espacement moindre entre deux arbres) mais tout en gardant une certaine organisation spatiale.

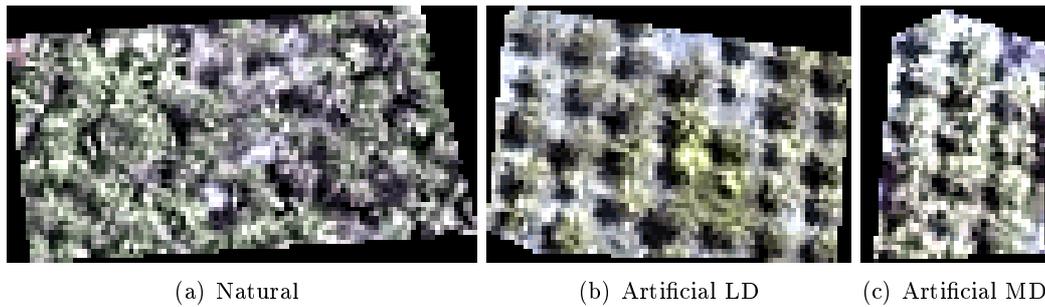


FIG. 10.1 – Exemple des différents types plantations.

La mesure de la qualité de la classification est obtenue en calculant la séparabilité des classes grâce à la distance de Jeffries-Matusita (équation (9.1)).

	Natural	Artificial LD	Artificial MD
Natural	0		
Artificial LD	1.44641	0	
Artificial MD	0.74986	0.40689	0

TAB. 10.1 – Séparabilité des classes.

Comme le montre le tableau 10.1, l'utilisation des informations spectrales seules ne permettent d'obtenir une classification correcte. En effet, nous voyons que la meilleur séparabilité ne dépasse pas 1,446 pour les plantations naturelles et elles sont encore plus faibles pour les plantations artificielles : 0,749 pour la classe Artificial MD et 0,40 pour la classe Artificial LD. Ceci est illustré par la figure 10.2 où sont présentées les résultats médiocres de la classification bayésienne et par SVM. Il est donc nécessaire d'ajouter de l'information dans les ensembles d'apprentissages. Aussi nous nous proposons d'ajouter aux informations spectrales l'utilisation d'informations texturales calculées à partir des descripteurs de texture d'Haralick qui prennent en compte les informations spatiales et spectrales.

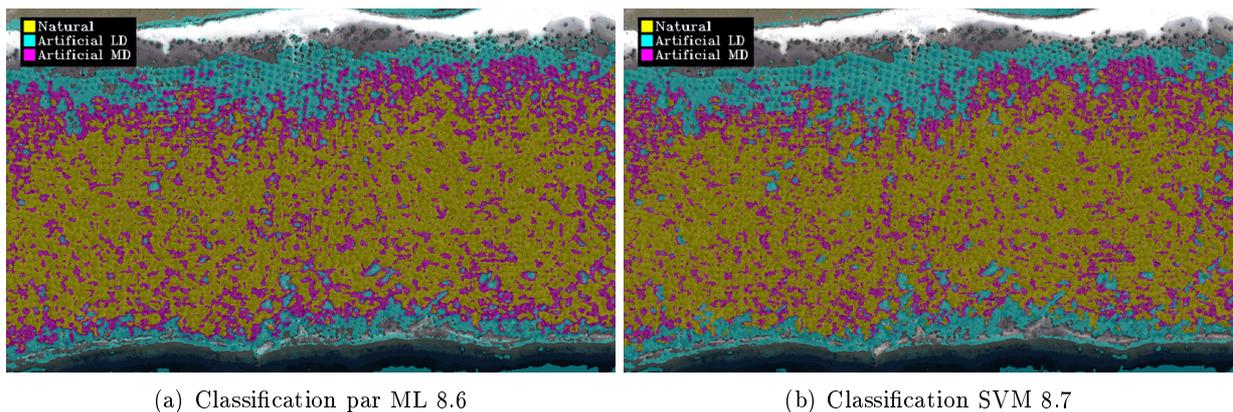


FIG. 10.2 – Résultat de la classification bayésienne et SVM avec uniquement les informations spectrales.

10.1.2 Choix de l'information spectrale

Une ACP est réalisée sur l'image RGB et on travaille sur la première composante qui contient la grande partie de l'information spectrale (voir section 5.2.3). Huit descripteurs de texture sont ensuite calculés sur cette composante. Il s'agit de la moyenne, la variance, l'homogénéité, le contraste, la dissimilarité, l'entropie, le Second Moment Angulaire, la corrélation. Pour ne privilégier aucune direction, chaque descripteur est calculé dans les quatre orientations (0° , 45° , 90° et 135°) et un descripteur

moyen est calculé à partir de ces quatre valeurs. Ensuite les informations de texture et les informations spectrales rassemblées et normalisées pour former une image à onze canaux (trois pour les informations spectrales RGB et huit pour les informations texturales).

Le calcul des informations de texture doit se faire dans une fenêtre glissante de taille optimale pour permettre de caractériser correctement chaque type de plantation. Une étude de la taille de la fenêtre est donc réalisée pour déterminer la taille optimale au sens de la meilleure séparabilité des classes. La figure 10.3 montre l'évolution de la moyenne totale des séparabilités des classes en fonction de la taille de la fenêtre utilisée pour le calcul des descripteurs de texture. Nous observons que la séparabilité est supérieure à 1,9 (indiquant une bonne prédiction de la classification) à partir d'une fenêtre glissante supérieure ou égale à 29×29 pixels. Nous choisissons donc la taille de la fenêtre égale à 29×29 pixels pour le reste de l'étude.

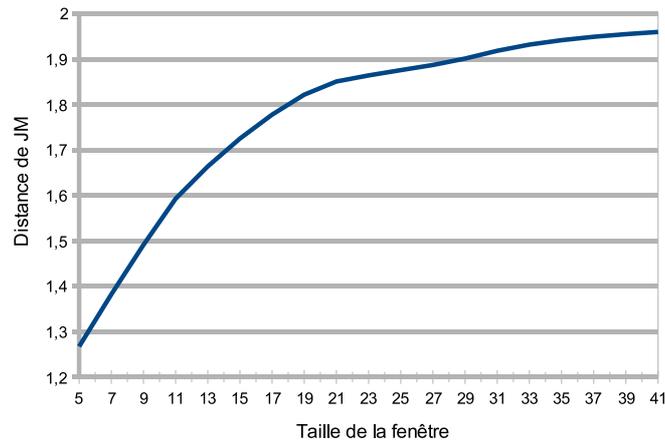


FIG. 10.3 – Évolution de la séparabilité des classes en fonction de la taille de la fenêtre des descripteurs de texture.

10.2 Choix du classifieur

10.2.1 Analyse des performances

Pour étudier la performance de chaque classifieur, nous réalisons une série d'expériences pour lesquelles nous faisons à nouveau varier la taille de la fenêtre glissante (de 5×5 à 41×41 pixels). Pour mesurer la performance de chaque classification, nous générons aléatoirement 3000 pixels dans chaque région d'apprentissage sur lesquels sont calculées les matrices de confusion. Le performance moyenne intra-classes est calculée par moyenne des éléments diagonaux de la matrice de confusion. Le tableau 10.2 montre un exemple de calcul de la performance moyenne.

Classes	Données de référence			Performance moyenne
	Natural	Artificial LD	Artificial MD	
Natural	98,63	0	14,33	89,22
Artificial LD	0	99,89	16,53	
Artificial MD	1,37	0,11	69,13	

TAB. 10.2 – Méthode de calcul de la performance moyenne intra-classe à partir de la matrice de cohérence.

La figure 10.4 présente le graphe des différentes performances des classifieurs en fonction de la taille de la fenêtre utilisée pour le calcul des descripteurs de texture. Nous voyons clairement que les classifications par SVM ou bayésienne donnent les performances les meilleures. En considérant une taille de fenêtre glissante de 5×5 pixels à 29×29 pixels, nous observons que ces deux méthodes ont des

performances équivalentes. Néanmoins, pour des fenêtres de taille supérieure à 29×29 , le classifieur SVM augmente ses performances alors que les performances du classifieur bayésien stagnent.

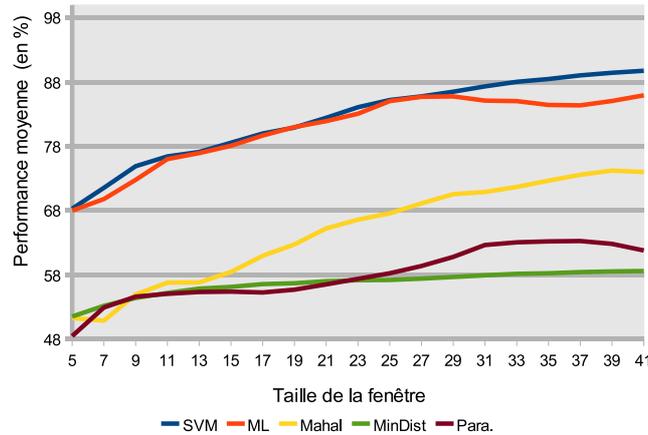


FIG. 10.4 – Performance des classifieurs.

Les trois autres classifieurs présentent des performances plus faibles :

- la classification par Distance de Mahalanobis a tendance à être meilleure jusqu'à atteindre un maximum pour une taille de fenêtre glissante égale à 37×37 pixels avec une valeur de 74.54%. Ceci permet tout de même de mettre en évidence l'influence des informations texturales sur la performance du classifieurs avec un gain de 24% ;
- en ce qui concerne la classification par Parallélépipède, la courbe des performances croît jusqu'à atteindre un maximum pour une taille de fenêtre de 37×37 pixels et un taux de pixels correctement classifiés ;
- la classification par Distance Minimale a les performances les plus faibles. La performance croît marginalement avec la taille de la fenêtre glissante. Les informations de texture ne permettent d'obtenir qu'un gain de 7% sur les performances de cette méthode de classification.

Cette étude met donc en évidence trois groupes de méthodes de classification en fonction de leurs performances pour la caractérisation des types de plantation :

- meilleures performance : SVM et ML,
- performances moyennes : Distance de Mahalanobis,
- performances faibles : Distance Minimale et Parallélépipède.

10.3 Application aux données Ikonos

Dans cette partie, nous utilisons le classifieur SVM pour caractériser les types de plantation de cocotiers. Dans un premier temps, nous cherchons à estimer les paramètres du classifieur qui obtiennent la meilleure classification. Puis nous appliquons la méthode sur la zone d'étude.

10.3.1 Choix des paramètres du classifieur

Les informations spectrales et texturales sont combinées comme vu dans la partie précédente 10.1.2 et constituent les données à classifier. Afin de mettre en évidence l'influence de chaque paramètre du classifieur dans les performances de la classification, nous fixons la taille de la fenêtre glissante à 37×37 pixels. Nous faisons varier les paramètres P et γ et mesurons la performance moyenne du résultat de la classification. Le paramètre P varie de 100 à 2000 avec un pas de 100. Quant au paramètre γ , il varie de 0,091 (qui est sa valeur par défaut, c'est-à-dire $\frac{1}{N}$ avec N le nombre de canaux) à 2,091 par pas de 0,1. la figure 10.5 montre l'influence de ces paramètre sur la qualité de la classification. La performance la plus faible est obtenue pour $P = 2000$ et $\gamma = 2,091$ avec un taux de classification correct de 86,961%. La meilleure classification est obtenue avec $P = 100$ et $\gamma = 0,991$ avec une performance moyenne de 89,899%. Pour une valeur de γ fixée à 0,991, les variations des valeurs de P ne permettent

pas d'améliorer la qualité de la classification (voir la figure 10.6(a)). En effet, à mesure que P croît, les performances décroissent en passant d'une performance moyenne de 89,899% à 88,842%. Cette tendance est générale comme le montre la figure 10.5. Le paramètre γ agit différemment sur la qualité de la classification. Comme le montre la figure 10.6(b) où le paramètre P est fixé à 100, la courbe croît jusqu'à atteindre son maximum pour une valeur de γ égale à 0,991 puis décroît.

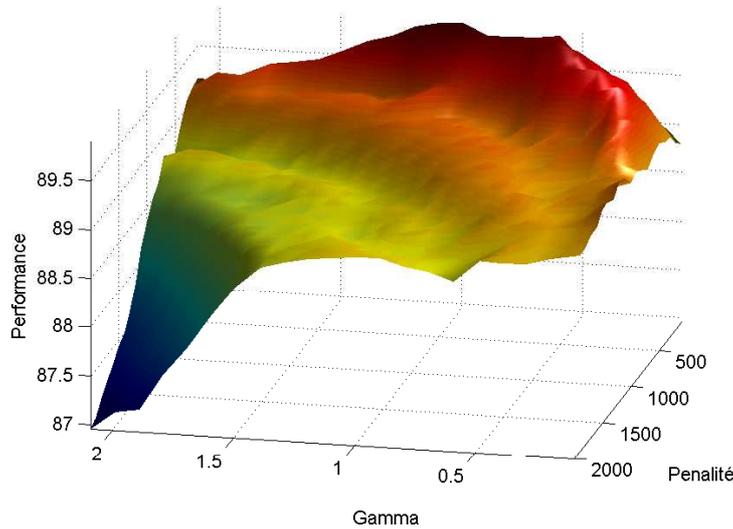


FIG. 10.5 – Influence des paramètres P et γ sur la performance du classifieur SVM.

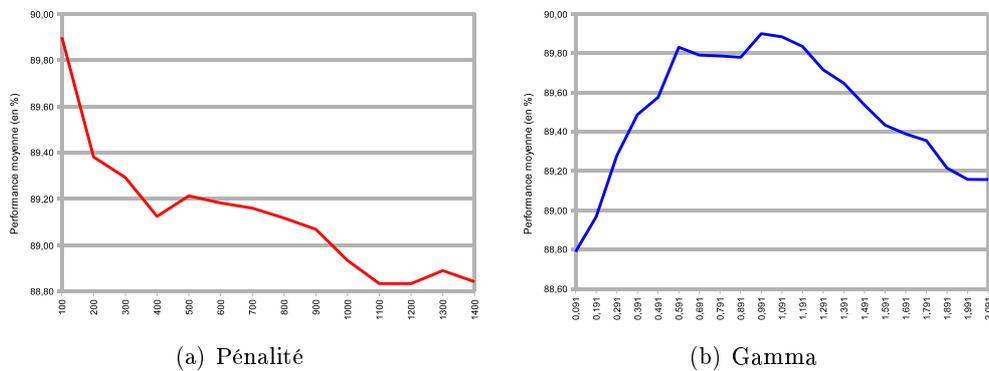


FIG. 10.6 – Influence des paramètres P et γ sur la performance du classifieur SVM.

10.3.2 Résultats

Nous fixons le paramètre de pénalité P à 100 et le paramètre γ à 0,991. La figure 10.7 montre le résultat de cette classification sur la zone d'étude située au Nord-Est de l'atoll de Vahitahi. Nous observons que les résultats sont très satisfaisants, visuellement parlant. Dans la classe Natural, en jaune, on distingue de temps en temps un alignement et une orientation. Mais l'ajout d'une nouvelle classe indiquant une cocoteraie artificielle à haute densité n'a pas apporté d'amélioration, bien au contraire, nous avons une dégradation de la qualité de la classification. En effet, avec le codage utilisé (informations spectrales et informations de texture), la séparabilité entre cette nouvelle classe et la classe Natural n'était pas suffisamment importante pour discriminer correctement ces deux classes. Nous avons donc fait le choix de ne pas l'utiliser.

Nous avons effectué le même travail sur l'atoll de Tikehau (voir figure 10.8) dont la résolution du pixel au sol est de 1 m.

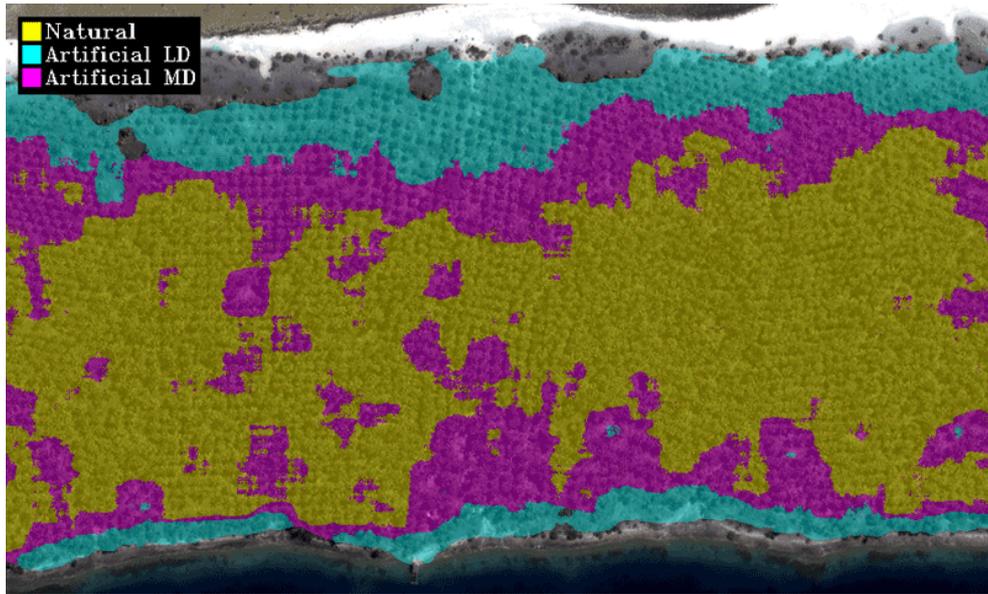


FIG. 10.7 – Résultat de la classification par SVM avec $P = 100$ et $\gamma = 0,991$.

Nous réalisons une classification de l'occupation des sols et des types de plantation sur l'ensemble de l'atoll à l'aide de la classification par SVM. Nous devons tout d'abord réaliser un jeu de classes définies à partir d'une version modifiée de la vérité terrain utilisée par [Jacq, 2006] et fournie par le cadastre. Les données n'étaient pas projetées de la même manière d'où la nécessité de recalibrer la vérité terrain sur notre image dans un premier temps. Ensuite, pour chaque classe, nous tirons aléatoirement 1100 pixels afin de définir notre ensemble d'apprentissage tels que précisé dans [Richards and Jia, 1999]. Le tableau 10.3 présente les classes utilisées ainsi que leur couleur associée.

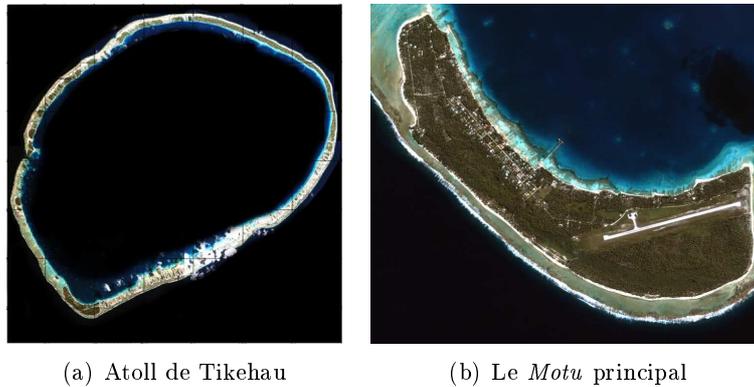


FIG. 10.8 – L'atoll de Tikehau et le Motu principal.

Vg B	Végétation basse	1
Vg H	Végétation haute	2
Vg E	Végétation épars	3
AF	Cocoteraie abandonnée	4
MF	Cocoteraie entretenue	5
UMF	Cocoteraie non entretenue	6
SW	Marécage	7
N-Veg	Pas de végétation	8

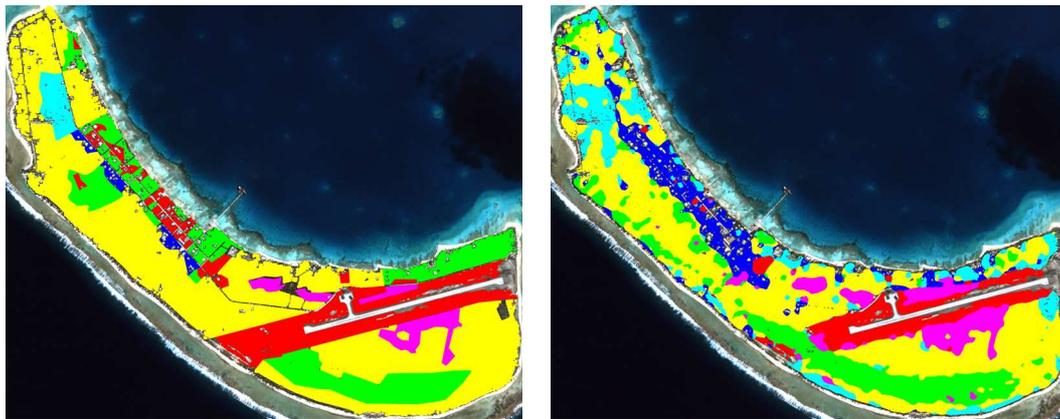
TAB. 10.3 – Acronyme des classes utilisées.

En utilisant la méthode décrite dans ce chapitre, c'est-à-dire choisir correctement les informations spectrales et texturales, nous calculons la séparabilité de chacune des classes qui est présentée dans le tableau 10.4. Sur le *motu* principal, *Haamea*, la classe **AF** n'est pas représentée mais elle est bien présente sur le reste de l'atoll. De plus, toutes les régions où la végétation n'est pas présente sont masquées afin de ne s'intéresser qu'à la typologie des plantations et éviter les erreurs de classification.

Classes	Vg B	Vg H	Vg E	AF	Vg MF	UMF	SW	N-Veg
Vg B	0	1,9629	1,9491	Type non présent sur le motu	1,9370	1,9602	1,8397	Zone sans végétation masquée
Vg H	1,9629	0	1,6466		1,3245	1,6475	1,9945	
Vg E	1,9491	1,6466	0		1,8557	1,8770	1,9960	
MF	1,9370	1,3245	1,8557		0	1,2258	1,9674	
UMF	1,9602	1,6475	1,8770		1,2258	0	1,9720	
SW	1,8397	1,9945	1,9960		1,9674	1,9720	0	
N-Veg								

TAB. 10.4 – Séparabilité des classes.

La figure 10.9 présente la vérité terrain utilisée ainsi que le résultat de la classification des types de plantation de cocotiers avec les SVM. À partir de cette classification, nous allons valider la méthode de segmentation des arbres et extrapoler le nombre d'arbres pour se rapprocher le plus possible de la réalité.



(a) Vérité terrain.

(b) Résultat de la classification.

FIG. 10.9 – La vérité terrain fournie par le cadastre et le résultat de la classification avec les SVM des types de plantations.

10.3.2.1 Validation et extrapolation du nombre de cocotiers

Nous allons effectuer une nouvelle validation de la méthode de dénombrement des cocotiers vu dans le chapitre 5 et quantifier les taux de détection en fonction de la topologie de la cocoteraie. Pour cela, nous nous intéressons à l'atoll de Tikehau sur lequel des missions terrain ont été effectuées dans le cadre de cette thèse (en été 2006) et par [Jacq, 2006]. Nous avons à notre disposition la base de données contenant toutes ses informations recueillies sur le terrain qui a été constituée pendant cette mission. Les mesures sont faites dans des placettes circulaires disposées en quinconce sur l'ensemble de l'atoll et d'un rayon de 12 m. Ces placettes sont repérées par leur position GPS. La figure J.1 présente une fiche terrain type pour chaque placette traitée.

De cette base, nous extrayons des informations telles que le numéro de la placette, les coordonnées de la placette, le nom du *motu*, le nombre d'arbres à l'intérieur de la placette, le nombre d'arbre dont

la hauteur est supérieure à 4 m et le nombre d'arbres dont la hauteur est supérieure à 5 m. Nous ne pouvons pas prendre uniquement le nombre d'arbres total dans une placette comme mesure de vérité terrain car cette mission date de 2006 alors que l'image à notre disposition est de 2003. C'est donc pour cela que nous avons extrait d'autres informations en rajoutant des contraintes sur la hauteur afin d'obtenir le nombre d'arbres le plus susceptibles d'être visible sur l'image lors de l'acquisition de cette dernière. Le tableau 10.5 présente une partie des données extraites de cette base auquel nous avons rajouté une colonne indiquant le nombre d'arbres détectés par notre méthode dans la placette.

ID placette	X	Y	Motu	Nb 12m	Nb H > 4m	Nb H > 5m	Méthode
0	366918,53	8327705,40	Haamea	7	7	5	8
1	368115,00	8328008,00	Haamea	14	14	14	7
2	367547,00	8327726,00	Haamea	14	10	10	6
3	367818,53	8327705,40	Haamea	15	13	12	7
4	368118,53	8327705,40	Haamea	14	4	4	7
10	368115,00	8328008,00	Haamea	10	10	9	7
11	366346,00	8328290,00	Haamea	17	16	15	9
12	366607,00	8328272,00	Haamea	9	5	5	7
⋮							

TAB. 10.5 – Exemple de données extraites de la base des relevés terrains.

En fusionnant les résultats de la classification par SVM des types de plantation (voir la figure 10.9(b)) avec la carte de densité calculée à partir de la position des arbres détectés (voir la figure 10.10(a)) et les mesures réalisées sur le terrain (voir la figure 10.10(b) où les placettes sont matérialisées par des cercles jaunes), nous pouvons déterminer le taux de détection en fonction du type de plantation.

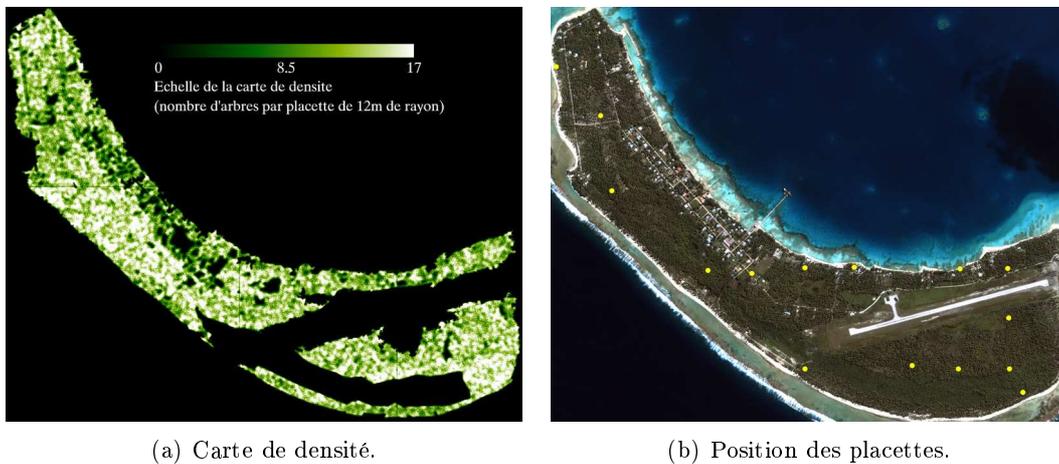


FIG. 10.10 – La carte de densité calculée à partir des positions des arbres détectés et la position des placettes dans lesquelles les mesures terrains sont réalisées.

Le tableau 10.6 résume les différents taux de détection en fonction de chaque classe mais aussi en fonction de la contrainte sur la hauteur lors des requêtes sur la base. Nous observons que notre méthode détecte entre 66% et 89% (pour les arbres dont la hauteur H est supérieure à 5 m). Enfin, à partir de toutes ces données, nous calculons un taux de détection moyen relatif à chaque classe. Nous pouvons maintenant extrapoler le nombre d'arbres en tenant en compte du taux de détection moyen relatif à chaque classe pour obtenir une estimation du nombre de cocotiers qui se rapproche le plus de la réalité. Le tableau 10.7 présente les résultats obtenus sur l'ensemble de l'atoll de Tikehau.

En annexe J, nous présentons dans le tableau J.1 les résultats détaillés pour chacun des *motu* qui composent l'atoll de Tikehau.

Classe	Nombre d'arbres				Performance			
	Nb 12 m	H > 4 m	H > 5 m	Méthode	% Diff 4 m	% Diff 5 m	% Diff Nb 12 m	% Moyen
Vg B	73	50	48	51	102,00%	106,25%	69,86%	92,70%
Vg H	228	206	185	132	64,08%	71,35%	57,89%	64,44%
Vg E	Pas de mesures terrains disponibles.							
AF	687	584	540	399	68,32%	73,89%	58,08%	66,76%
MF	174	133	123	102	76,69%	82,93%	58,62%	72,75%
UMF	8	6	6	4	66,67%	66,67%	50,00%	61,11%
SW	22	21	16	11	52,38%	68,75%	50,00%	57,04%
N-Veg	Pas de mesures terrains disponibles.							

TAB. 10.6 – Taux de détection en fonction des classes.

	Classe								Total
	Vg B	Vg H	Vg E	AF	MF	UMF	SW	N-Veg	
Taux moyen	92,70%	64,44%	0,00%	66,76%	72,75%	61,11%	57,04%	0,00%	
DéTECTÉS	13771	21735	650	87824	23439	1334	973	904	150630
EXTRAPOLÉS	14855	33728	650	131546	32220	2183	1706	904	217792

TAB. 10.7 – Extrapolation finale du nombre d'arbres sur l'ensemble de l'atoll de Tikehau.

10.4 Conclusion

Nous avons montré qu'utiliser uniquement les informations spectrales n'était pas suffisant pour caractériser les types de plantation de cocotiers, mais que cette tâche pouvait être accomplie grâce aux informations de texture qui prennent en compte la disposition spatiale des objets.

L'étude comparative des classifieurs a mis en évidence que le classifieur SVM par noyau RBF (fonctions à base radiales) et le classifieur bayésien se distinguent des autres par des performances bien meilleures. Cette classification permet aussi d'obtenir une estimation du nombre d'arbre dans chaque classe puisqu'on connaît explicitement le nombre de cocotiers calculé par la partie 5.2. Bien que la démarche donne de bons résultats, elle possède l'inconvénient d'être lente. En effet, plus la taille de l'image augmente, plus le temps nécessaire au calcul des informations de texture est long et de plus la classification est faite pixel par pixel. Sur le *motu* principal, dont l'image fait 3355×2920 pixels, le temps nécessaire pour calculer les descripteurs de texture avec une fenêtre glissante de 23×23 pixels, pour une direction donnée, est d'environ 13 minutes. Il faut donc ensuite comptabiliser le temps nécessaire au calcul dans les trois autres directions. La classification SVM nécessite quant à elle plusieurs heures de calcul. Un moyen d'accélérer la méthode serait de réaliser en parallèle des classifications sur les informations spectrales et une des informations de texture, puis de fusionner les résultats. Nous proposons dans le chapitre suivant une approche pour caractériser les plantations de cocotiers en analysant la répartition spatiale des cocotiers.

Caractérisation des cocoteraies par l'analyse de la distribution spatiale des cocotiers

Bien que la caractérisation des différents types de plantations soit réalisable grâce à la classification par SVM, cette méthode présente le principal inconvénient d'être gourmande en ressources informatiques. Nous proposons une méthode alternative en utilisant uniquement les coordonnées des arbres puisque ces dernières peuvent être déduite du processus de segmentation décrit dans la partie II. Nous introduisons la notion de processus ponctuel de Poisson qui permet de caractériser les plantations sauvages. Un procédé local d'estimation du paramètre du processus ponctuel de Poisson est proposé pour rendre compte des configurations diverses trouvées dans les images.

11.1 Analyse spatiale des coordonnées

11.1.1 Processus ponctuel

Définition 11.1 *On appelle processus ponctuel une variable aléatoire dont les réalisations sont des ensemble de points du plan. L'ensemble des réalisations est appelé semis (de points).*

Un processus ponctuel peut générer une infinité de semis de points différents tout en gardant des propriétés communes. Inversement, pour un semis de points donné, il existe une infinité de processus ponctuels dont ce semis pourrait être une réalisation.

Pour l'analyse spatial d'un semis de points donné, il est nécessaire de faire l'hypothèse qu'il existe un processus ponctuel dont le semis de points a une forte probabilité d'en être une réalisation. De plus, la seule mesure de la réalisation à disposition (c'est-à-dire le semis de points observé) impose d'autres hypothèses telles que l'homogénéité, l'isotropie et l'ergodicité pour rendre possible l'analyse de la structure spatiale.

Définition 11.2 *Un processus ponctuel est dit homogène (ou stationnaire) s'il est invariant par translation, c'est-à-dire si ses propriétés ne varient pas d'une position à l'autre dans l'espace.*

Définition 11.3 *Un processus ponctuel est dit isotrope s'il est invariant par rotation, i.e que ses propriétés ne varient pas avec l'orientation dans l'espace.*

11.1.2 Processus ponctuel de Poisson

Un processus ponctuel de Poisson uniforme (appelé également *Complete Spatial Randomness*) dans le plan de paramètre $\lambda > 0$ vérifient les propriétés suivantes :

- le nombre de point à l'intérieur de toute région B suit une loi de Poisson d'espérance $\lambda\mathcal{A}(B)$,
- étant donné n points dans une région, ces points y sont uniformément distribués,
- les contenus de deux sous-régions disjointes sont indépendants.

Définition 11.4 Soit $X = \{X_i\}$ une famille de variables aléatoires à valeurs dans $S \subset \mathbb{R}^2$. Cette famille est appelé Processus Ponctuel de Poisson (PPP) si :

1. pour tout $B \subset S$, le nombre de point appartenant à B , noté $\text{Card}(X \cap B)$, suit une loi de Poisson de paramètre $\lambda \mathcal{A}(B)$:

$$\mathbb{P}(\text{Card}(X \cap B) = n) = e^{-\lambda \mathcal{A}(B)} \frac{(\lambda \mathcal{A}(B))^n}{n!} \quad (11.1)$$

où $\mathcal{A}(B)$ est l'aire de B .

2. si B_1 et B_2 , deux sous-ensembles de S , sont disjoints, alors $\text{Card}(X \cap B_1)$ et $\text{Card}(X \cap B_2)$ sont indépendants.

Le paramètre λ , appelé intensité, caractérise un processus de Poisson. Pour un processus homogène, λ est constante et peut être calculé par l'estimateur du maximum de vraisemblance qui est $\hat{\lambda} = \frac{N}{D}$ où N est le nombre de points et D la surface. La probabilité qu'un point du semis soit présent dans une surface élémentaire dS centrée en (x, y) est :

$$\mathbb{P}(\text{un point dans } dS) = \lambda dS \quad (11.2)$$

Lorsque ce paramètre dépend des coordonnées spatiales, ceci parce que la distribution diffère d'une région à une autre, le processus est appelé processus ponctuel de Poisson spatial (*Spatial Poisson Point Process* ou SPPP). Un SPPP peut être utilisé pour caractériser une distribution aléatoire de points dans l'espace en effectuant le test statistique suivant : test de l'hypothèse \mathcal{H}_0 que le processus est poissonnien contre celle d'un processus non poissonnien. Sous l'hypothèse \mathcal{H}_0 , le semis de point est un processus ponctuel de Poisson et donc que le nombre de points dans toute une région B suit une loi de Poisson d'intensité $\lambda \mathcal{A}(B)$. Sous hypothèse \mathcal{H}_0 , l'espérance du nombre de point dans une région B est donc :

$$\mathbb{E}(\text{Card}(X \cap B)) = \lambda \mathcal{A}(B)$$

Supposons maintenant que l'intensité dépende de la localisation spatiale x , on peut montrer que :

$$\mathbb{E}(\text{Card}(X \cap B)) = \int_B \lambda(x) dx \quad (11.3)$$

La figure 11.1 illustre des simulations de processus ponctuel de Poisson homogène avec différentes valeurs d'intensité. La figure 11.2 illustre des simulations de processus ponctuels de Poisson non-homogènes, c'est-à-dire avec une l'intensité λ non constante et dépendante de la localisation spatiale.

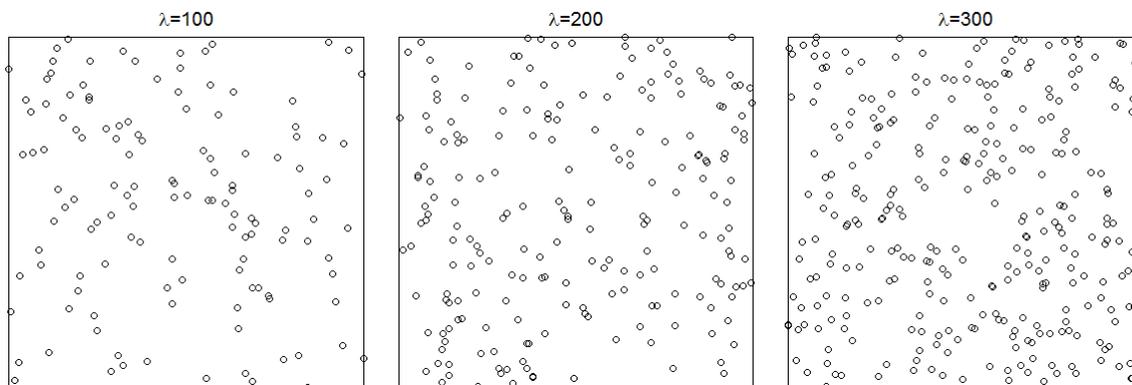


FIG. 11.1 – Simulation de processus ponctuel de Poisson homogène dans un carré unitaire avec $\lambda = 100, 200$ and 300 .

Si l'intensité λ n'est pas spatialement homogène, on peut alors essayer de l'approximer localement.

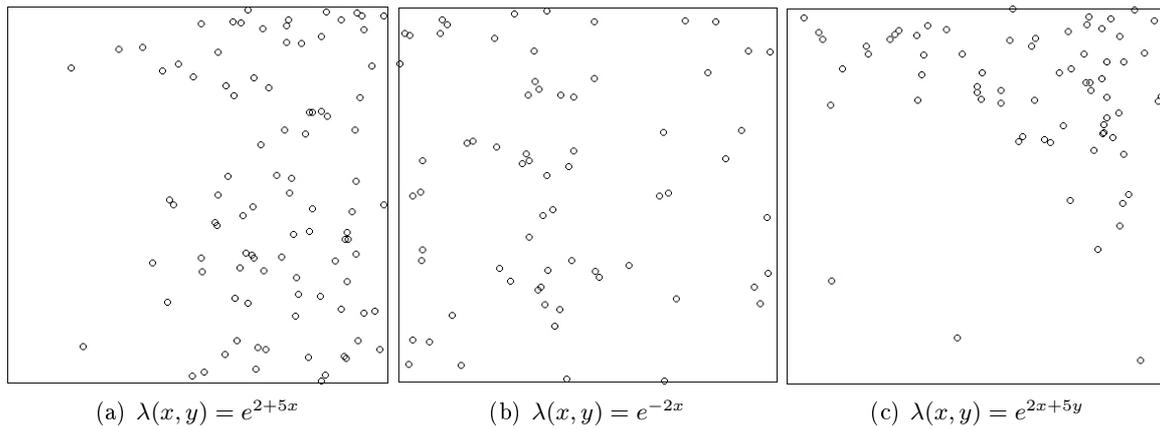


FIG. 11.2 – Simulation de processus ponctuel de Poisson non-homogène dans un carré unitaire avec différentes fonction d'intensité $\lambda(x, y)$.

11.1.3 Estimation de l'intensité par un SPPP

Nous étudions ici trois méthodes de la littérature [Ripley, 1981, Diggle, 1983, Cressie, 1993] qui permettent d'estimer les lois de processus ponctuel de Poisson :

- analyse par Quadrants,
- estimation de la fonction d'intensité λ par des méthodes de noyaux,
- analyse du second ordre.

Méthode des Quadrants

La méthode des Quadrants consiste à subdiviser l'espace en sous-régions B_1, B_2, \dots, B_n de taille égale appelées Quadrants. Les points à l'intérieur de chaque quadrant sont comptés : $n_j = \text{Card}(X \cap B_j)$ pour $j = 1, \dots, m$. Sous hypothèse \mathcal{H}_0 , c'est-à-dire si le nombre de points dans une région suit une loi de Poisson, les $\{n_j\}$ sont des variables aléatoires indépendantes et identiquement distribuées de valeur égale.

Pour chacun des quadrants, la valeur d'intensité λ est estimée, notée $\hat{\lambda}$. Afin de vérifier leur homogénéité spatiale, un ensemble de réalisations sont simulées à partir de cette fonction d'intensité $\hat{\lambda}$ et ces simulations sont comparées à la réalisation à estimer au moyen de diverses mesures telles que :

1. test statistique testant si le semis de point est un PPP [Pearson, 1900] ;
2. la distance de Clark Evans (CE) [Clark and Evans, 1954] ;
3. indice de dispersion (ID) [Ludwig and Reynolds, 1988, Pielou, 1977, Jayaraman, 2000] : indice calculé à partir du nombre d'événements dans un Quadrant et le nombre total de Quadrants ;
4. la distance des indices de dispersion (I) [Johnson and Zimmer, 1985] : basée sur la distance d'un point à un événement.

Les méthodes basées sur les quadrant sont dépendantes du partitionnement de l'espace et ne mettent pas en évidence les variations spatiales de la fonction d'intensité λ à l'intérieur des régions mais uniquement entre deux régions adjacentes.

Méthode des noyaux

Une autre approche consiste à estimer la fonction d'intensité λ en utilisant des méthodes de noyaux [Diggle, 1985]. Il s'agit d'estimer la fonction d'intensité λ à l'aide d'une fenêtre glissante en comptabilisant les événements présents dans cette fenêtre et centrée sur chacun des points du semis. Cette carte d'intensité est ensuite convoluée avec un noyau gaussien. Cette méthode permet de visualiser les disparités spatiales de la fonction d'intensité λ à différentes valeurs de l'écart type du noyau gaussien telle que elle est illustrée par la figure 11.3, c'est-à-dire à différentes échelles spatiales dans l'image. Il s'agit donc d'une approche multi-échelles.

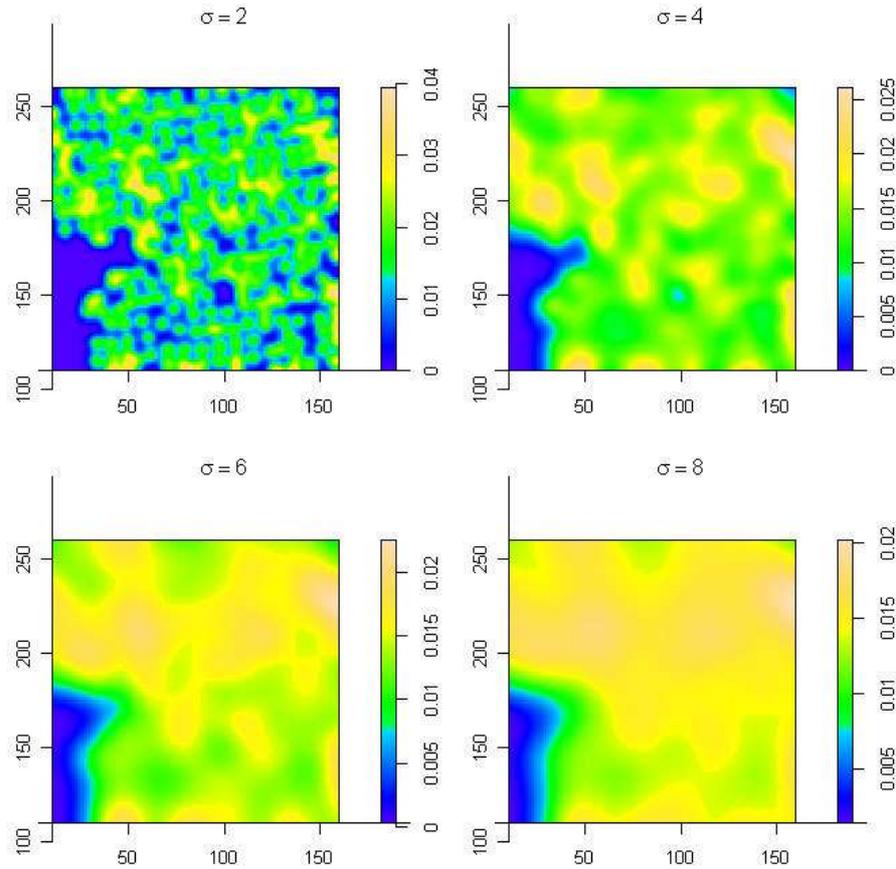


FIG. 11.3 – Estimations de la fonction d'intensité $\lambda(u)$ avec plusieurs valeurs de l'écart type du noyau de convolution Gaussien.

Analyse de second ordre

La fonction de Ripley [Ripley, 1977, Ripley, 1981], notée K , permet de réaliser une analyse de second ordre d'un semis de points. Cette fonction analyse toutes les distances point à point, à la différence de l'analyse de premier ordre telle que la plus courte distance. La fonction K d'un processus ponctuel est définie telle que $\lambda K(r)$ soit l'espérance du nombre de voisins à distance r d'un point quelconque du semis. Sous hypothèse \mathcal{H}_0 , l'espérance du nombre de points dans un cercle centré en x et de rayon r est $\lambda\pi r^2$.

La forme générale de cette fonction K est la suivante :

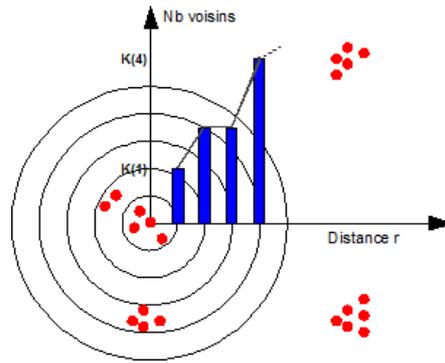
$$K(r) = \frac{1}{\lambda^2 \mathcal{A}(S)} \sum_{i=1}^n \sum_{j \neq i}^n \mathbb{1}_{\{\|x_i - x_j\| \leq r\}} e(x_i, x_j, r) \quad (11.4)$$

avec $e(x_i, x_j, r)$ une fonction corrigeant le biais sur les bords du domaine d'étude.

On pourra lire [Baddeley, 1998] pour avoir un exemple pratique d'une telle fonction. S est le domaine d'étude et $\mathbb{1}$ est une fonction indicatrice, c'est-à-dire que l'on a :

$$\mathbb{1}_{\{\|x_i - x_j\| \leq r\}} = \begin{cases} 1 & \text{si } \|x_i - x_j\| \leq r \\ 0 & \text{sinon} \end{cases} \quad (11.5)$$

Sous hypothèse \mathcal{H}_0 , nous avons $K(r) = \pi r^2$, les valeurs telles que $K(r) < \pi r^2$ indique un semis aggloméré alors que $K(r) > \pi r^2$ suggère un semis régulier. La figure 11.4 montre la méthode de calcul de la fonction $K(r)$.


 FIG. 11.4 – Méthode de calcul de la fonction $K(r)$.

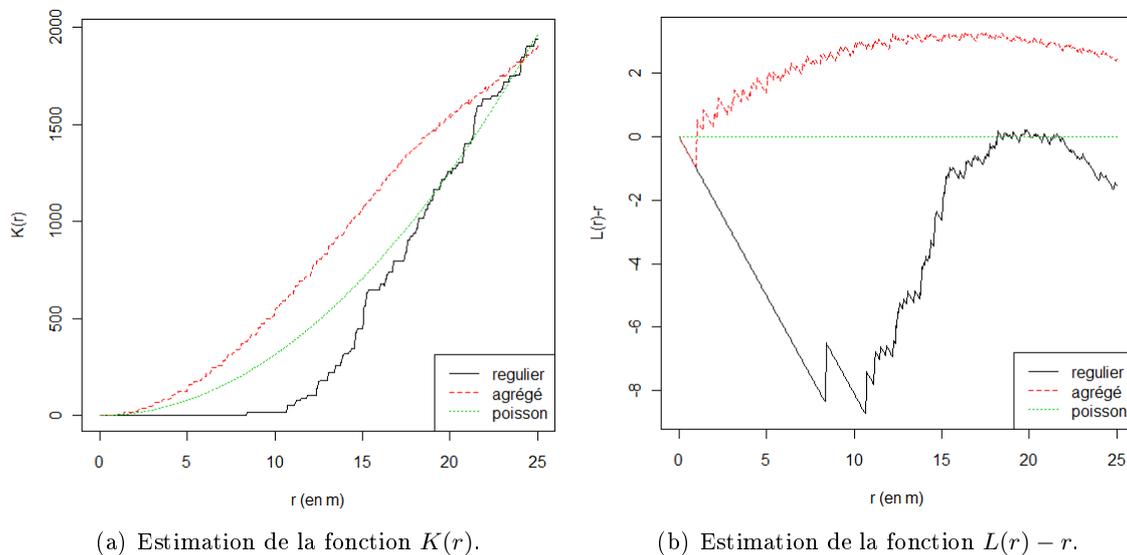
L'interprétation de la fonction K peut prêter à confusion. Une transformation communément utilisée est la fonction L [Besag, 1977] définie telle que :

$$L(r) = \sqrt{\frac{K(r)}{\pi}} \quad (11.6)$$

Sous hypothèse \mathcal{H}_0 , $L_{pois}(r) = r$. La fonction K est une fonction cumulative et il devient difficile d'interpréter correctement les interactions spatiales pour de grandes distances. La fonction de corrélation de paires de points $g(r)$ [Stoyan and Stoyan, 1994, Stoyan and Stoyan, 1996] s'y prête mieux. Elle est définie par :

$$g(r) = \frac{1}{2\pi r} \frac{\partial K(r)}{\partial r} \quad (11.7)$$

Sous hypothèse \mathcal{H}_0 , $g(r) = 1$. Alors que des valeurs telles que $g(r) < 1$ suggère une régularité entre les points et des valeurs telles que $g(r) > 1$ suggère une agglomération.


 (a) Estimation de la fonction $K(r)$.

 (b) Estimation de la fonction $L(r) - r$.

 FIG. 11.5 – Estimation des fonctions K et L pour différents types de semis : régulier, agrégé et poissonnien.

11.2 Application à la caractérisation de la cocoteraie

Grâce aux méthodes d'analyse spatiale de semis de points, nous pouvons donc caractériser des répartitions spatiales dans les cocoteraies. À cette fin, nous définissons deux régions, notées région #01 et région #04 dans le *motu* principal de Tikehau (voir la figure 11.6).



FIG. 11.6 – L'atoll de Tikehau et les zones d'études sur le *motu* principal.

La première étape de cette analyse est de segmenter les arbres, d'en déduire leur centre et d'extraire leur coordonnée spatiale. Pour chacune de ces deux régions, nous définissons une sous-région correspondant à un carré de 150 m de côté (soit une surface de 2,25 Ha). Nous appelons sous-région #01 la sous-région définie dans la région #01 et nous faisons de même pour la région #04 (voir la figure 11.7). Dans la sous-région #01, la distance minimale entre les arbres est de 3,6 m et elle est de 4 m dans la sous-région #04. La figure 11.7 montre l'ensemble du semis de point pour les régions #01 et #04 et leur sous-région correspondante. Le tableau 11.1 résume les statistiques pour ces régions (la surface, le nombre d'arbre et l'intensité moyenne).

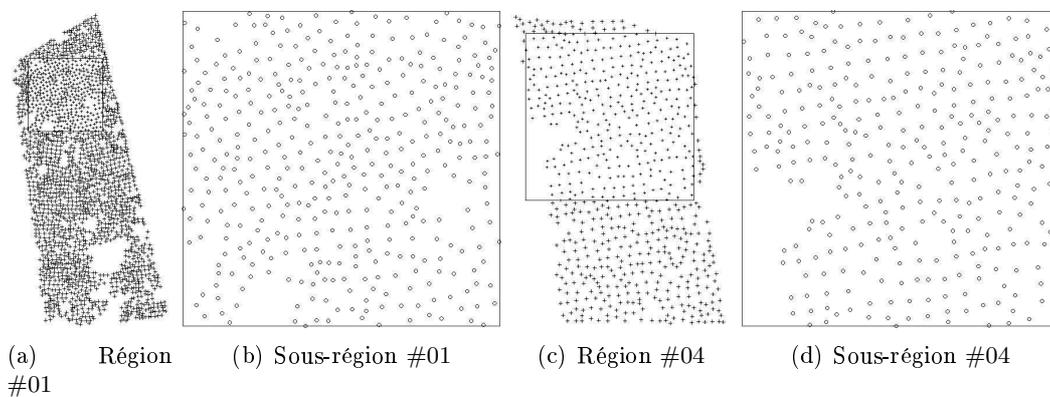


FIG. 11.7 – Les régions étudiées ainsi que leur sous-régions.

	Région #01	Sous-région #01	Région #04	Sous-région #04
Surface (en Ha)	12.69	2.25	3.99	2.25
Nombre d'arbres	1950	399	632	316
$\lambda(\text{arbres}/m^2)$	0.0154	0.0177	0.0158	0.014

TAB. 11.1 – Quelques statistiques calculées sur les régions et sous-régions d'études.

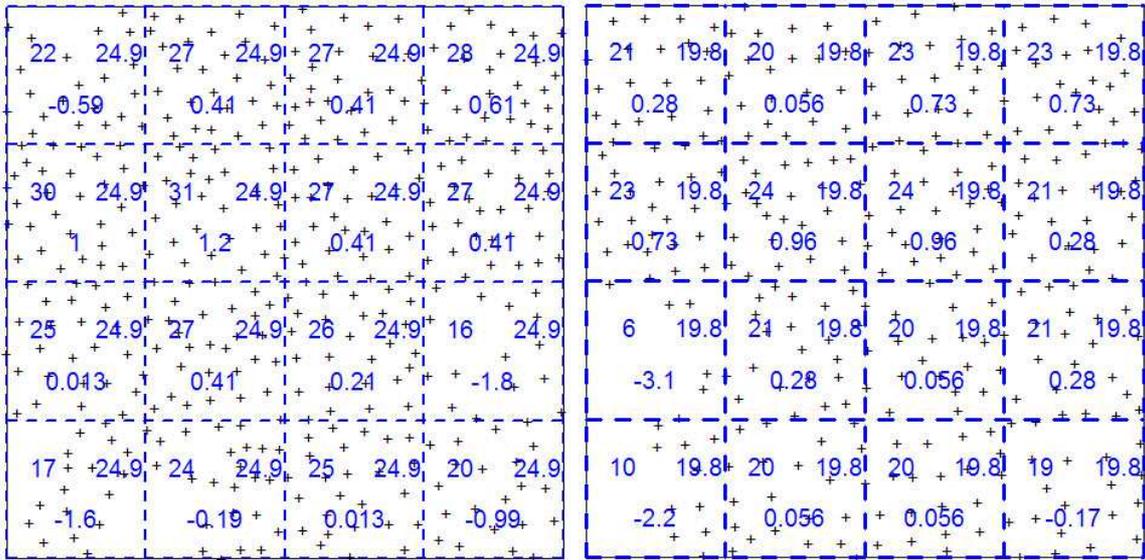
La seconde étape de l'analyse spatiale est d'estimer la fonction de densité λ . Pour cela, nous utilisons les méthodes décrites dans la section 11.1.3.

11.2.1 Estimation de la fonction d'intensité $\lambda(x)$

Dans un premier temps, l'analyse des quadrants permet de visualiser la variabilité de la fonction d'intensité $\lambda(x)$. Les sous-régions #01 et #04 sont partitionnées en 16 régions de taille égale. La figure

11.8 montre le partitionnement des sous-régions en surface de taille égale avec diverses informations pour chaque quadrant tels que le nombre d'arbres observé (en haut à gauche), le nombre attendu (en haut à droite) et le test des résidus de Pearson défini tel que :

$$\text{Test de Pearson} = \frac{\text{Card}(\text{observation}) - \text{Card}(\text{attendu})}{\sqrt{\text{Card}(\text{attendu})}} \quad (11.8)$$



(a) Sous-région #01

(b) Sous-région #04

FIG. 11.8 – Test statistique de Pearson pour les deux sous-régions.

Le tableau 11.2 montre les différentes valeurs des distances calculées afin de déterminer si le semis de points rejette l'hypothèse \mathcal{H}_0 .

	Sous-région #01			Sous-région #04		
	CE	ID	I	CE	ID	I
Valeur de l'indice	1.30	3.81	2.53	1.27	5.58	5.95
Test Statistiques	8.76	3.81	2.53	7.26	5.58	25.20
Rejète \mathcal{H}_0	oui	non	oui	oui	oui	oui

TAB. 11.2 – Détermination des indices de dispersion.

Finalement, la méthode des quadrants peut donner une bonne indication sur la fonction d'intensité $\lambda(x)$ pour chacune des surfaces du partitionnement. Mais les résultats dépendent beaucoup de la manière dont ce partitionnement est réalisé.

Afin d'obtenir une meilleure visualisation de la variabilité spatiale de la fonction d'intensité $\lambda(\vec{x})$, nous utilisons la méthode des noyaux. Pour les calculs, nous avons utilisé un écart type $\sigma = 12 m$ pour la convolution gaussienne. Comme le montre la figure 11.9, les semis de points des sous-régions #01 et #04 ne sont pas homogènes, c'est-à-dire que la fonction d'intensité λ n'est pas constante dans tout l'espace mais que celle-ci présente des variations liées à la position dans l'espace étudié.

Pour caractériser plus finement les plantations de cocotiers, nous utilisons l'analyse de second ordre appliqué à la position de des cocotiers.

11.2.2 Analyse de second ordre de la cocoteraie

À partir de la fonction K (équation (11.4)), nous calculons la fonction L (équation (11.6)) associée à chacune des sous-régions #01 et #04. La figure 11.10 montre les résultats des calculs où $L(r) - r$ est graphé en fonction de r . Sous hypothèse \mathcal{H}_0 , nous avons $L(r) - r = 0$ (courbe en pointillés rouge

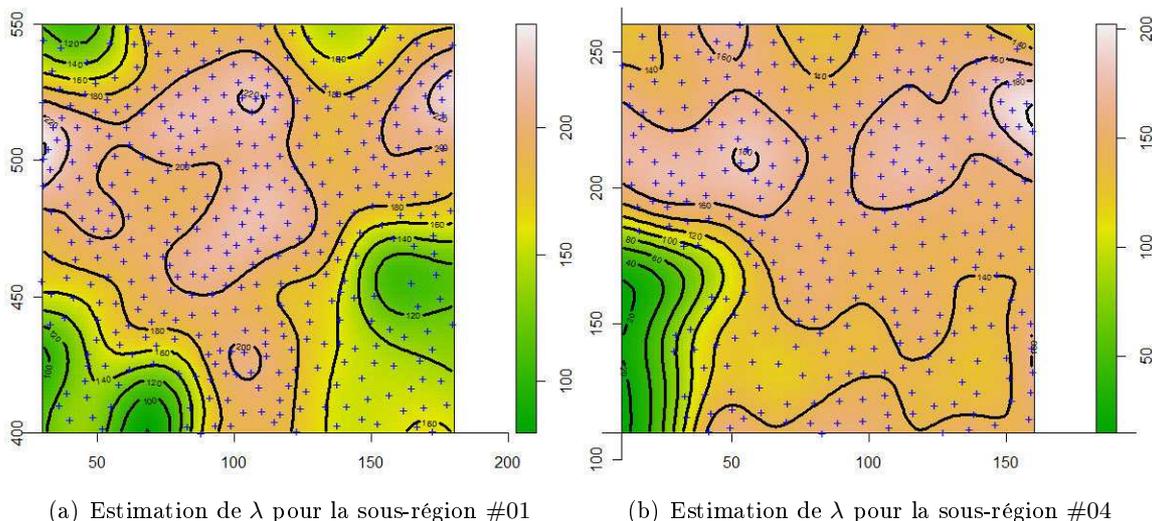


FIG. 11.9 – Estimation des fonctions d'intensité λ pour les sous-régions #01 et #04. Les densités sont exprimées en nombre d'arbres par hectare.

sur les figures). Si le semis de point présente des agrégations, alors la courbe $L(r) - r$ tend à devenir positive et si les points présentent une régularité alors la courbe $L(r) - r$ tend à être négative. Les intervalles de confiance sont obtenus en réalisant 1000 simulations et en prenant la courbe maximale et la courbe minimale.

Les points des deux sous-régions #01 et #04 montrent une régularité significative pour une distance $r < 9,5 m$ (la fonction L calculée avec ces observations est sous la borne inférieure de l'intervalle de confiance à 95%) avec une distance minimale entre deux arbres de l'ordre de $3,5 m$ pour la sous-région #01 et de l'ordre de $4 m$ pour la sous-région #04. Nous observons aussi une régularité entre les points pour des distances $r \in]11 m; 15,5 m]$ dans la sous-région #01. Finalement, le semis de point de la sous-région #01 suit l'hypothèse \mathcal{H}_0 pour des distances $r \in [9 m; 11 m[$ et $r > 15,5 m$.

Par contre, les points du semis de la sous-région #04 montrent une agglomération pour $r \in [31 m; 32 m]$. Enfin, pour des distances $r \in]9,5 m; 31 m]$, le semis suit l'hypothèse \mathcal{H}_0 car la courbe correspondante à notre observation se trouve entre les intervalles de confiance (haute et basse).

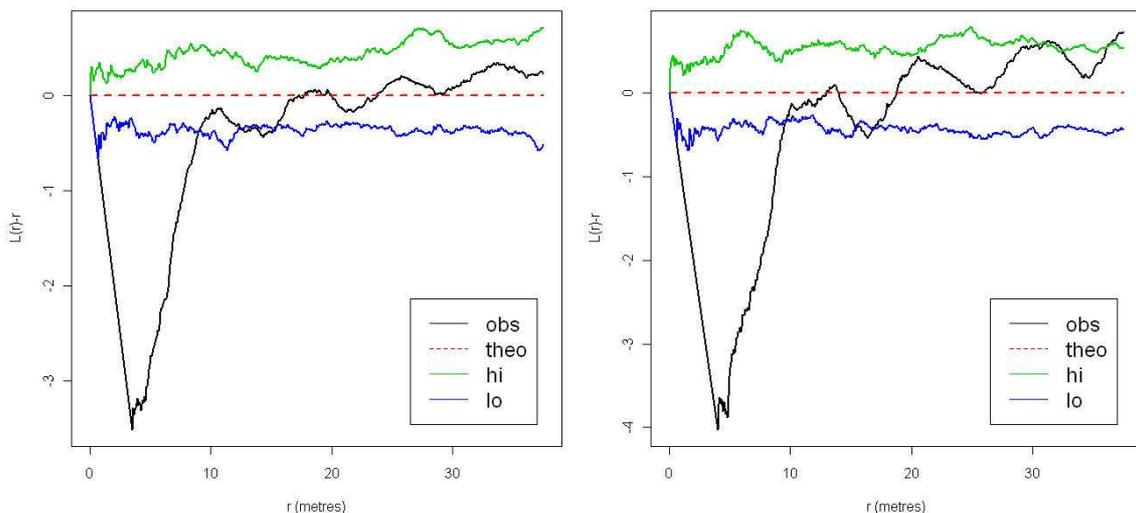


FIG. 11.10 – Graphe de la fonction $r \mapsto L(r) - r$, les intervalles de confiance à 95% (haute en vert, basse en bleu), la valeur théorique de la courbe (ligne en pointillés rouge) et l'observation (ligne noire).

La fonction de corrélation g de paires de points (équation (11.7)) est calculée pour chacune des deux sous-régions. La figure 11.11 montre cette fonction avec les deux intervalles de confiance (la borne supérieure en vert et la borne inférieure en bleu), la valeur théorique pour un processus ponctuel de poisson uniforme (ligne en pointillé rouge) et la courbe correspondant à notre observation (en noir). Sous hypothèse \mathcal{H}_0 , $g(r) = 1$, les valeurs telles que $g(r) < 1$ suggèrent une régularité entre les points et les valeurs telle que $g(r) > 1$ suggèrent plutôt une agrégation des points.

Avec cette fonction g , nous pouvons observer que les arbres de la sous-région #01 sont agglomérés pour des distances $r = [6 \text{ m}; 10 \text{ m}]$, $[15 \text{ m}; 17 \text{ m}]$, $[23 \text{ m}]$, $[24 \text{ m}; 25 \text{ m}]$, et que ceux de la sous-région #04 le sont aussi pour des distances $r = [7 \text{ m}; 10, 5 \text{ m}]$, $[17 \text{ m}; 21 \text{ m}]$, $[26; 29 \text{ m}]$.

En comparant les figures 11.10 et 11.11, nous concluons que les deux fonctions L et g mettent en évidence les mêmes type de distribution des arbres mais à différentes échelles (c'est-à-dire à des valeurs de distance différentes).

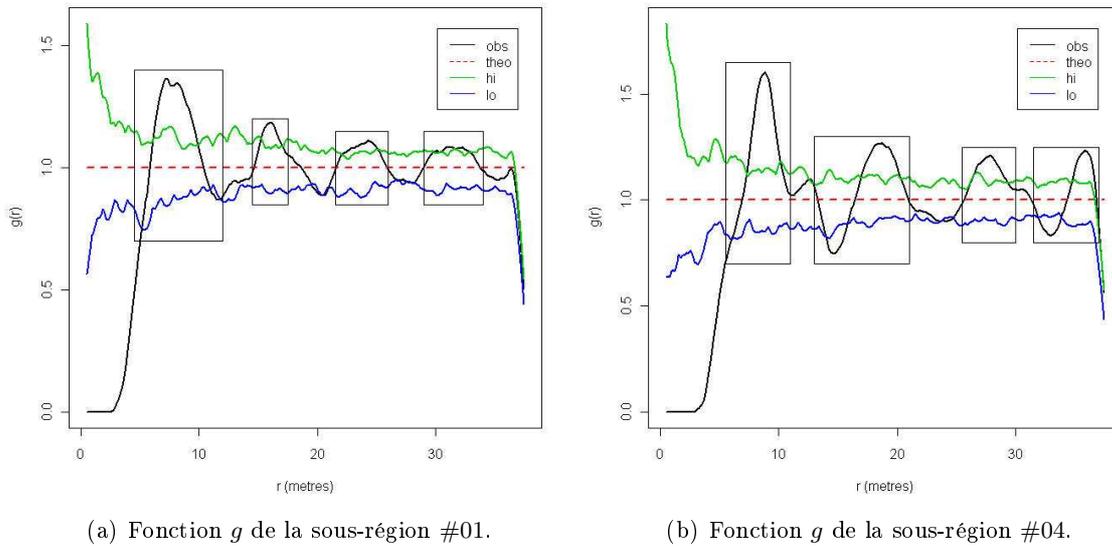


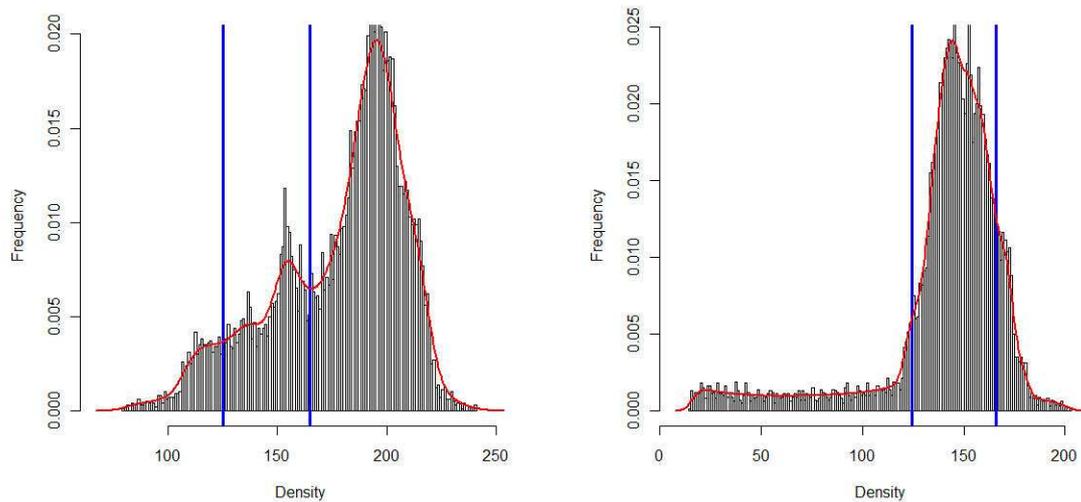
FIG. 11.11 – Fonctions de corrélation de paire de points (fonction g) avec la borne supérieure (ligne verte) et la borne inférieure (ligne bleue) de l'intervalle de confiance, la courbe théorique (ligne en pointillés rouge) et celle de notre observation (ligne noire).

11.2.3 Classification des plantations de cocotiers

On souhaite maintenant caractériser chaque région obtenue avec la méthode des noyaux. L'espèce de cocotiers largement répandue et utilisée dans les cocoteraies est le *Grand Local*. Pour une exploitation et un rendement maximal de la plantation, il est recommandé d'avoir un espacement minimal de $8,5\text{m}$ entre chaque arbre (planté suivant un motif en carré ou en quinconce), ce qui implique une densité à l'hectare de l'ordre de 140 à 150 arbres [Chantry, 1983]. En prenant en considération ces recommandations, nous examinons l'histogramme des densités (ou la carte des fonctions d'intensité) et nous définissons trois classes prenant en compte la densité théorique (nombre d'arbres à l'hectare) ainsi que l'écart type σ_{KD} du résultat de l'estimation de la fonction d'intensité grâce à la méthode des noyaux. Ces trois classes sont définies telles que :

$$\left\{ \begin{array}{ll} \text{seuil} < 140 - \sigma_{KD} & \text{plantation moins sauvage : classe 1} \\ 140 - \sigma_{KD} \leq \text{seuil} \leq 150 + \sigma_{KD} & \text{plantation entretenue : classe 2} \\ \text{seuil} > 150 + \sigma_{KD} & \text{plantation sauvage ou naturelle : classe 3} \end{array} \right. \quad (11.9)$$

La figure 11.12 montre les histogrammes des fonctions d'intensité estimées pour les deux sous-régions #01 et #04. Les valeurs des seuils haut et bas sont matérialisées par des lignes verticales bleues. Pour la sous-région #01, ce sont 125 arbres/hectare et 165 arbres/hectares. Et pour la sous-région #04 elles sont de 124 arbres/hectare et 166 arbres/hectare.

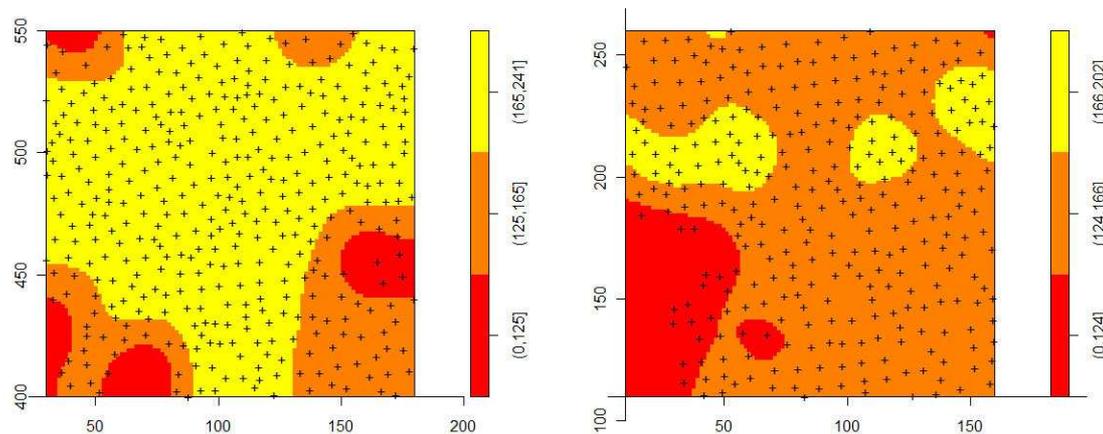


(a) Histogramme de la carte de densité de la sous-région #01.

(b) Histogramme de la carte de densité de la sous-région #04.

FIG. 11.12 – Les histogramme des densités estimées pour les sous-régions #01 et #04. Les lignes verticales bleues représentent les valeurs de seuil.

La figure 11.13 présente les résultats obtenus du seuillage en trois classes des cartes de densités avec les positions des arbres superposées.



(a) Seuillage de la carte de densité de la sous-région #01.

(b) Seuillage de la carte de densité de la sous-région #04.

FIG. 11.13 – Seuillages des sous-régions #01 et #04.

Le tableau 11.3 liste quelques statistiques pour chacune des trois classes obtenues par seuillage de l'histogramme dont les seuils sont calculés selon les règles données dans l'équation (11.9). Bien que la surface et le nombre d'arbres des deux sous-régions #01 et #04 présents dans la classe 3 (représentant les plantations sauvages ou naturelles) soient différents, nous observons une densité moyenne égale. Nous observons la même chose concernant la classe 2 (représentant les plantations maintenues). Les résultats pour la classe 1 dans la sous-région #01 montre que seuls sept arbres sont présents. Bien que l'intensité moyenne soit proche de celle de la classe 1 de la sous-région #04, le nombre d'arbres n'est pas suffisamment important pour conclure. Nous concluons que les plantations entretenues et sauvages sont les deux principaux types de plantation présentes dans la sous-région #01.

Afin de prendre des décisions sur les propriétés spatiales de chacune des régions segmentées, nous analysons les diagrammes Quantile-Quantile des résidus [Baddeley et al., 2005]. Ce diagramme est un outil de diagnostic qui permet de comparer graphiquement si un échantillon suit une loi de distribution particulière. Nous concluons qu'une distribution suit une loi de distribution donnée si l'analyse des

	Sous-région #01			Sous-région #04		
	Classe 1	Classe 2	Classe 3	Classe 1	Classe 2	Classe 3
Nombre d'arbres	7	78	314	20	243	53
Intensité moyenne	0.00414	0.0154	0.02	0.00631	0.0148	0.0185
Surface (en m^2)	1692	5080.5	15727.5	3170.5	16465.5	2864.25

TAB. 11.3 – Résumé de chacune des zones segmentées pour les sous-régions #01 et #04.

résidus montre que la courbe reste dans l'intervalle de confiance.

Dans notre étude, nous prenons comme loi de référence un processus ponctuel de Poisson et nous vérifions si les distributions spatiales dans chacune des classes suivent ou non un tel processus. La figure 11.14 présente le diagramme Quantile-Quantile des résidus pour chacune des trois classes de la sous-région #04. Pour chaque diagramme, l'intervalle de confiance est dessiné en pointillé rouge et le modèle choisit comme référence est en noir. les figures 11.14(a) et 11.14(c) suggèrent qu'un modèle uniforme de Poisson est approprié pour les classes 1 et 3. Ceci permet de conclure que ces régions représentent des plantations sauvages où nous ne détectons aucune structure (alignement, répétition) spatiale entre les arbres. Par contre, étant donné que l'intensité moyenne de la classe 3 est supérieure à celle de la classe 1, les plantations de la classe 3 sont plus sauvages que celle représentée par la classe 1.

La figure 11.14(b) indique que le modèle de référence utilisé (processus ponctuel de Poisson uniforme) n'est pas approprié pour la classe 2 et donc celle-ci est caractérisée comme une plantation entretenue par l'homme.

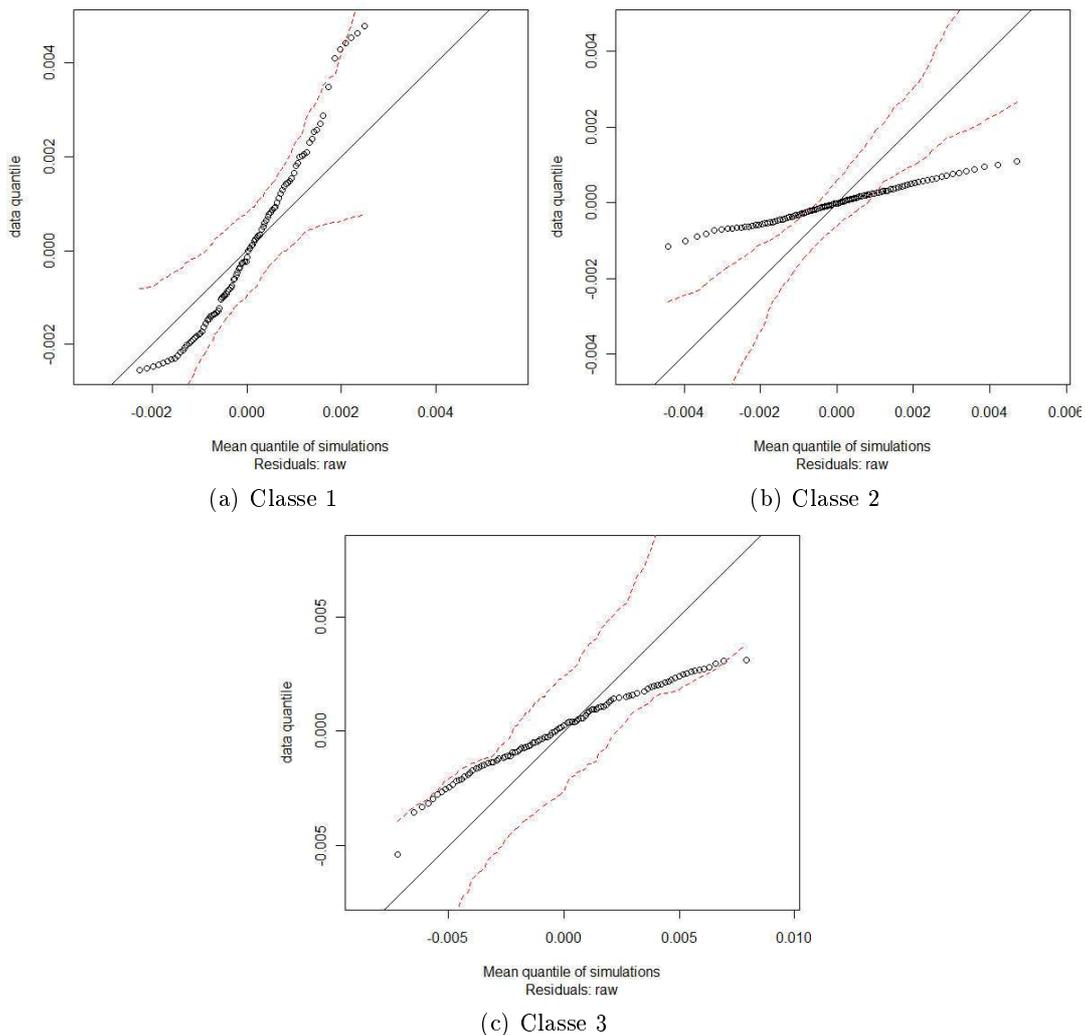


FIG. 11.14 – Diagramme Quantile-Quantile des classes segmentées de la sous -région #04.

Conclusion

Dans cette partie, nous avons montré que l'utilisation des techniques d'analyse spatiale permettent de caractériser une cocoteraie à partir de la position des cocotiers. La dernière partie a consisté à segmenter les régions en trois classes en analysant les histogrammes des densités et en seuillant la carte de d'intensité estimée par la méthode des noyaux. Le seuil est calculé en prenant en compte les valeurs théoriques de densité pour une pousse optimale du cocotier en fonction de l'espèce présente. L'analyse des régions segmentées a permis d'obtenir des informations sur la distribution spatiale des cocotiers et de confirmer ou pas la nature poissonnienne d'une région. Une perspective sera de calculer les seuils des histogrammes en utilisant des compositions de gaussienne ou la méthode d'Otsu pour améliorer la segmentation.

Conclusion et perspectives

Dans cette thèse, nous avons proposé une méthode de recensement des cocotiers et de caractérisation, en terme de différenciation des configurations spatiales, des cocoteraies dans les images Ikonos dépourvues de bande proche infrarouge. Les informations extraites du recensement permettent de définir un programme d'optimisation des ressources. L'exploitation de la cocoteraie est la seule source de revenu, en dehors du tourisme, pour les habitants des atolls où aucune activité de perliculture n'est implantée. La production de *coprah* est majoritairement exportée à l'étranger ce qui représente un intérêt économique non négligeable pour le Territoire de la Polynésie Française.

La thèse est d'abord axée sur les méthodes de détection des arbres dans des images satellites dépourvues de bande proche infrarouge. Sans cette bande proche infrarouge, qui permet normalement de calculer des indices discriminant de végétation, il a fallu trouver une méthode alternative pour générer les masques de la végétation haute. La génération de ces masques est une étape importante car elle permet de limiter les erreurs de détection. L'étape suivante est le processus de segmentation des arbres. Pour cela, nous effectuons une fusion des canaux spectraux par analyse en composantes principales et travaillons sur la première composante qui regroupe toutes les informations des trois canaux spectraux rouge, vert et bleu (avec une inertie de 99%). Puis, nous choisissons judicieusement des marqueurs dans l'image qui serviront de source d'immersion à la ligne de partage des eaux (LPE). Enfin, le résultat de la segmentation est affiné en éliminant les régions aberrantes. La segmentation permet, en plus d'obtenir un dénombrement des cocotiers, d'extraire des informations relatives à chaque arbre telles que :

- le centre de l'arbre, calculé comme étant le centre pondéré par informations spectrales,
- la surface de la couronne,
- la couleur moyenne de l'arbre pour chaque canal rouge, vert et bleu.

La méthode de recensement a été validée à l'aide deux protocoles :

- la première validation est réalisée par photo interprétation humaine. Un opérateur interprète les structures présentes dans l'image et marque celles qui correspondent à un cocotier. Les résultats obtenus sont comparés à ceux de la méthode,
- la seconde validation est réalisée par des mesures de densités effectuées sur le terrain. Ces mesures sont repérées par GPS et les valeurs obtenues sont mises en corrélation avec les cartes de densité calculées à partir de la position des arbres détectés.

La validation par photo interprétation a permis de mettre en évidence une erreur de détection de 10%. Une telle validation permet de conclure sur la pertinence de l'algorithme à partir des données images mais pas de la réalité sur le terrain. Une mission terrain a donc été nécessaire et elle nous a aussi permis de rendre compte de la topologie des cocotiers. Plusieurs difficultés, non visibles directement sur les images, ont été rencontrées lors de la mission terrain. La première est la zone de corail soulevé, appelé *FEO*, qui n'a pas permis de réaliser les mesures de densité dans cette région. La seconde est l'emplacement d'un cocotier par rapport aux cocotiers voisins. Il a fallu distinguer trois strates qui caractérisent les différentes configurations. La corrélation des mesures effectuées sur le terrain avec les résultats du dénombrement des cocotiers dans les images Ikonos a permis de calculer des taux de

détection pour la validation de la méthode sur des zones test. Un comptage exhaustif a été réalisé sur un *motu* et les résultats obtenus sont équivalents. Nous avons pu conclure que la méthode proposée pour le dénombrement est efficace avec un taux de détection moyen de 61%. Ce chiffre peut paraître peu élevé mais ceci est dû d'une part au fait que l'acquisition des images a été réalisée en 2003 et la mission terrain a été effectuée en été 2006, d'autre part que les acquisitions images ne peuvent donc pas observer tous les cocotiers présents effectivement sur le terrain.

Une fois la méthode validée, nous l'avons appliquée à l'ensemble des *motu* qui constituent l'atoll de Tikehau. Le passage à l'échelle a permis de mettre en évidence la robustesse de la méthode. Lors du traitement de ces *motu*, nous avons rencontrés plusieurs problèmes :

- la taille des images qui a imposé un découpage en imagerie,
- la gestion des zones d'ombres qui a dû être faite à part,
- et la perte d'arbres isolés occasionnée lors de l'application d'opérateurs morphologique pour la construction du masque de la végétation.

Une étude sur l'estimation de la hauteur des cocotiers a aussi été effectuée dans le cadre de cette thèse. L'analyse de l'ombre portée et de la configuration de la scène (position du soleil et du satellite) permettent de déduire la hauteur d'un arbre ou d'un bâti.

La seconde partie de la thèse est orientée vers la classification des sols et la caractérisation de la cocoteraie. Une étude des performances des méthodes de classification pour la classification des sols a d'abord été réalisée. Nous avons montré qu'il est nécessaire de coder différemment l'information d'un pixel par l'ajout des informations de texture aux informations spectrales afin de garantir la meilleure qualité possible de la classification. L'apport des informations de texture a permis de correctement classer la végétation haute par rapport à la végétation basse. Au terme des expérimentations effectuées, nous avons conclu que le classifieur SVM avec un noyau de fonctions à base radiale (RBF) se prête bien à la résolution d'un tel problème avec une erreur de classification de 4%. La méthode est appliquée à différentes images des autres atolls et les résultats obtenus ont permis de générer des cartes thématiques d'estimation des surfaces de cocoteraies (en ne considérant que celles d'une surface supérieure à 1 *Ha*). Ces cartes sont actuellement exploitées par le Service du Développement Rural (SDR) de la Polynésie Française.

Concernant la caractérisation de la cocoteraie en terme de répartition spatiale des cocotiers, deux approches ont été étudiées. La première approche consiste à classer directement les cocoteraies à partir des valeurs radiométriques de chaque pixel. Là encore, nous utilisons les informations de texture et les informations spectrales pour coder l'information d'un pixel. Une étude comparative des performances des classifieurs a été réalisée en comparant les résultats de la classification avec une vérité terrain fournie par le Service de l'Urbanisme de la Polynésie Française. Nous avons montré que le classifieur SVM avec un noyau RBF et le classifieur bayésien ont les meilleures performances. Grâce à cette classification, nous avons pu estimer le nombre de cocotiers dans chaque classe. Les missions terrains effectuées ont permis de déterminer un taux de détection relatif à chaque classe. Grâce à ce taux de détection, une extrapolation adaptative a pu être réalisée pour se rapprocher du nombre de cocotiers réellement présents. L'inconvénient de cette approche est sa lenteur car la classification est réalisée pixel par pixel et nécessite beaucoup de temps de calcul et de ressources informatiques.

La seconde approche consiste à caractériser les cocoteraies par l'analyse de la répartition spatiale des cocotiers grâce à leurs coordonnées. L'analyse d'un ensemble de points est faite grâce à diverses techniques. Ces techniques sont basées sur les processus ponctuels de poisson. Un processus ponctuel de poisson uniforme permet de caractériser une plantation naturelle, c'est-à-dire une plantation dans les cocotiers sont disposés aléatoirement. Grâce à ces outils, nous avons pu segmenter les régions d'étude en trois classes. L'étude de la répartition spatiale des cocotiers dans chaque région segmentée a permis d'obtenir des informations sur la distribution spatiale des cocotiers et de confirmer ou non la nature poissonnienne d'une région.

Le processus de dénombrement des cocotiers nécessite l'intervention de l'utilisateur pour la génération des masques (choix des zones d'apprentissages). L'acquisition de données dans le proche infrarouge permettrait de réduire cette supervision et d'automatiser la génération des masques de végétation. Ces

indices de végétation permettraient aussi d'éviter les erreurs de détection dues à la présence de sable dans les images. L'acquisition de données récentes est nécessaire pour mieux déterminer les erreurs de détection car il s'avère qu'une grande surface de la cocoteraie a été défrichée pour un projet de scierie dans le but d'exploiter le bois des arbres séniles. Bien que la vérité photo établisse une erreur de détection de l'ordre de 10%, nous avons des difficultés à séparer des arbres trop proches. Ceci peut être résolu en améliorant le calcul des marqueurs qui servent de source d'immersion de la LPE. Une approche possible est de réaliser des modèles en 3D mieux adaptés au cas des cocotiers que ceux vu dans la section 4.2. En particulier, il faut prendre en compte des troncs qui ne sont pas obligatoirement rectilignes (voir la fiche terrain dans la figure J.1).

Nous avons vu que, sur certaines images, l'ombre portée des nuages ne permet pas de classifier correctement la végétation. Dans cette thèse, nous avons traité ce cas séparément. Une détection des ombres grâce à des méthodes de seuillage permet d'obtenir la région correspondante à l'ombre portée d'un nuage, et nous avons réussi à dénombrer les cocotiers dans cette région. Une approche alternative serait de réaliser une correction radiométrique des zones d'ombres de manière à ce que la classification de la végétation n'échoue pas dans ces régions. La correction radiométrique consiste à restituer l'information des surfaces ombragées.

Une seconde perspective est d'améliorer la méthode de détection des nuages en utilisant leurs ombres portées. L'approche proposée serait la suivante :

1. détection des zones d'ombres à l'aide des méthodes proposées dans la littérature ou en recherchant une nouvelle plus adaptée. Ceci fournit l'initialisation de l'algorithme,
2. le nuage à segmenter est modélisé sous forme d'un contour actif qui peut alors se déformer de deux façons :
 - déplacement latéral, le long de l'axe ombre-soleil,
 - déformation locale du contour de façon à épouser au mieux les contours des nuages (voir la figure 12.1).

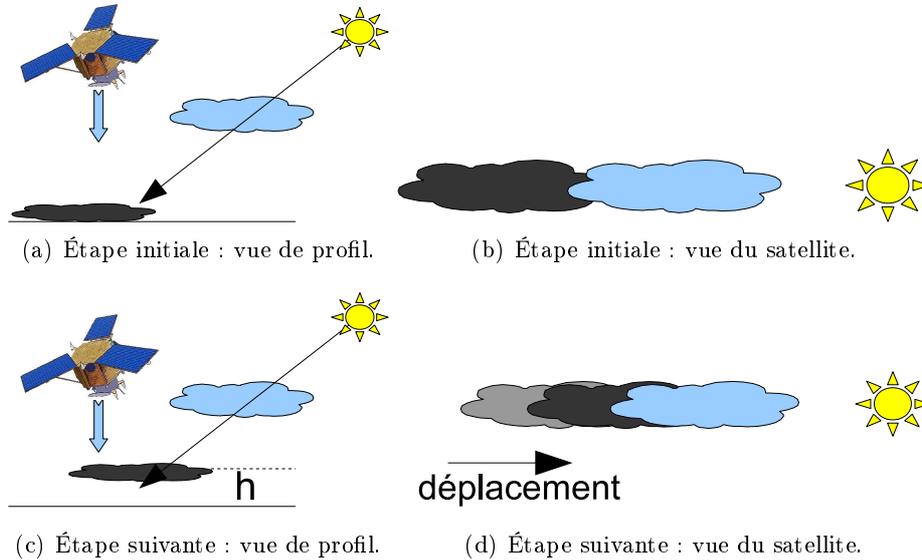


FIG. 12.1 – Processus de déplacement de l'ombre selon la direction opposée de la source de lumière.

Le déplacement du contour actif et sa déformation peuvent se faire soit simultanément soit séquentiellement. La position finale du contour actif le long de l'axe nuage-ombre donnera donc l'estimation de la hauteur de la base du nuage.

Nous avons vu que les ombres portées permettent d'estimer la hauteur des arbres. Cette méthode doit également détecter les cocotiers isolés et estimer leur hauteur. Grâce à l'ombre détectée, il est possible de rechercher, dans une direction privilégiée, l'arbre auquel elle appartient. Une approche contour actif de type *Level-sets* permettrait alors de gérer le problème de mise en correspondance entre les ombres et les cocotiers dans le cas où plusieurs cocotiers sont présents.

Quatrième partie

Annexes

Téledétection Optique

Cette annexe présente la télédétection optique, c'est-à-dire à l'acquisition d'images dans le spectre visible du rayonnement électromagnétique. Les caractéristiques du satellite Ikonos-2 et les systèmes embarqués qui servent à l'acquisition des images sont présentés.

A.1 Le satellite Ikonos

Ce satellite d'observation américain, d'une masse d'environ 700 kg, a été fabriqué par Lockheed Martin. Après le premier lancement raté du 29 avril 1999 d'Ikonos 1, Ikonos 2 (figure A.1) fut mis en orbite le 24 septembre 1999. Ce satellite a pour mission d'effectuer des prises de vue à très haute résolution, par exemple dans les zones urbaines.

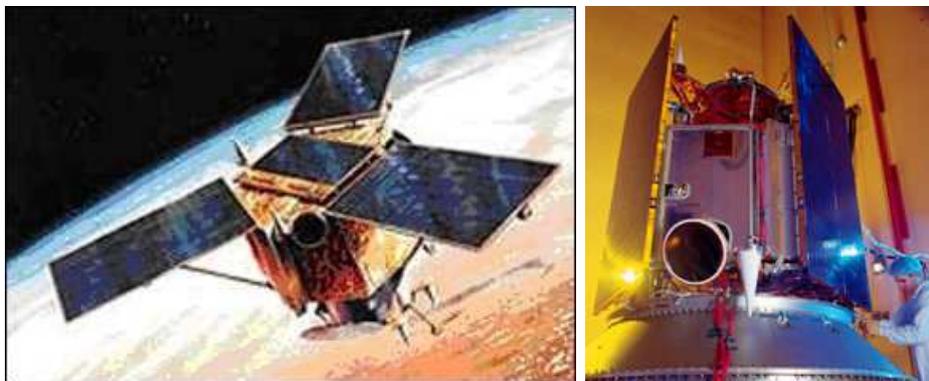


FIG. A.1 – Le satellite Ikonos-2.

A.2 Caractéristiques orbitales

Le satellite Ikonos 2 évolue à une altitude moyenne de 681 km, sur une orbite circulaire quasi polaire caractérisée par une inclinaison de $98,1^\circ$ (condition d'héliosynchronisme). La période de révolution est de 98 min. Ces caractéristiques orbitales ont pour conséquence de permettre l'acquisition des images avec une couverture complète entre les parallèles 81° nord et sud. La figure A.2 présente le tracé de l'orbite du satellite Ikonos-2. À chaque révolution, le satellite survole les régions de même latitude à l'ouest du passage précédent et la distance séparant deux traces consécutives (2700 km à l'équateur) est très supérieure à la largeur du champ d'observation du satellite (11 km). Pour pallier cette difficulté, le radiomètre équipant le satellite Ikonos peut être dépointé vers la gauche ou vers la droite de la direction de progression du satellite.

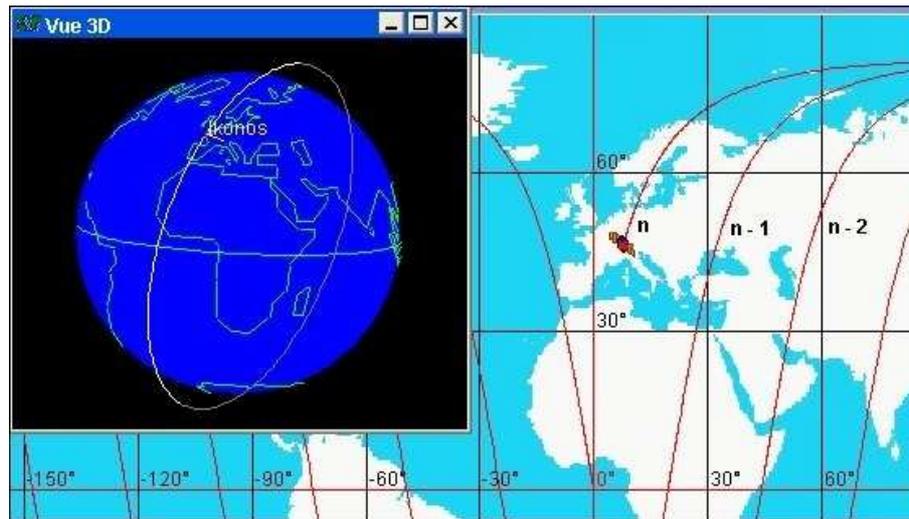


FIG. A.2 – Orbite du satellite Ikonos et sa trace au sol.

A.3 Systèmes embarqués

L'instrumentation embarquée est destinée à l'acquisition d'images à haute résolution.

Le radiomètre permet ainsi d'obtenir des images couvrant un champ d'observation de $11 \text{ km} \times 11 \text{ km}$, avec une résolution spatiale de 1 m en mode panchromatique et de 4 m en mode multispectral. En fait, au point sous-satellite (au nadir), la résolution est de 0,82 m en mode panchromatique et de 3,2 m en mode multispectral.

Toutefois, du fait de la faible largeur de la fauchée, un point particulier de la surface terrestre n'est survolé que tous les 35 jours, mais la possibilité de dépointage du capteur permet de ramener ce délai à 3 jours si besoin est. Le tableau A.1 présente les caractéristiques spectrales et la figure A.3 illustre les réponses spectrales relatives de chacune des bandes du satellite Ikonos 2 [GeoEye, 2006, Dial and Grodecki, 2003, Imaging, 2005, Cook et al., 2001].

mode	canal	bande spectrale	résolution spatiale	fauchée
Multispectral	1	0,445 – 0,516 nm (bleu)	4m	11 km
	2	0,506 – 0,595 nm (vert)	4m	11 km
	3	0,632 – 0,698 nm (rouge)	4m	11 km
	4	0,757 – 0,853 nm (proche ir)	4m	11 km
panchromatique		0,526 – 0,929 nm (vis et proche ir)	1m	11 km

TAB. A.1 – Caractéristiques principales du radiomètre équipant le satellite Ikonos 2.

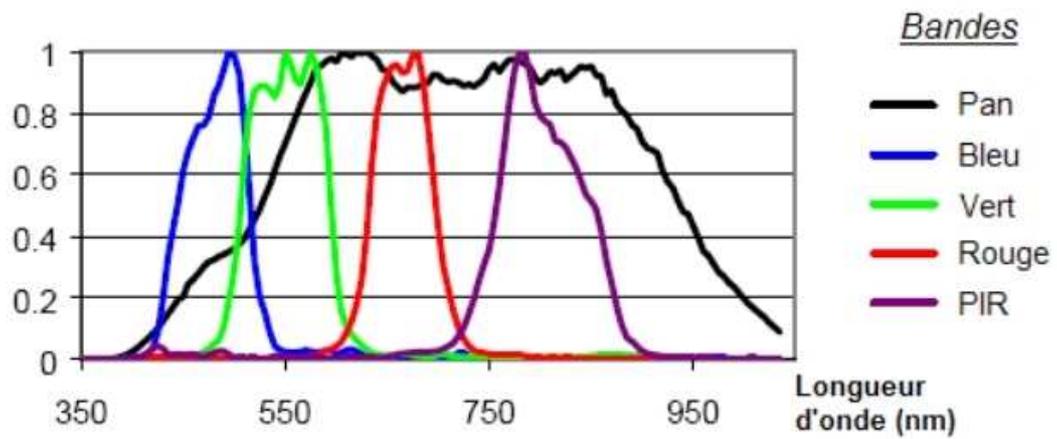


FIG. A.3 – Réponses Spectrales de la bande pan-chromatique et des bandes multi-spectrales.

Interaction avec la végétation

Cette annexe présente les caractéristiques spectrales de la végétation ainsi que les différents indices qui peuvent être calculés à partir des images satellites.

B.1 Caractéristiques spectrales de la végétation

Les propriétés optiques des couverts végétaux dépendent à la fois de la nature des plantes qui les composent, de leur état physiologique, des organes végétaux notamment les feuilles et leur arrangement spatial ainsi que du sol sous-jacent [Guyot, 1995].

B.1.1 Propriétés optiques des feuilles

Selon les types de végétation et la saison, les feuilles sont les principaux responsables de la diffusion et l'absorption de la lumière dans un couvert végétal. En particulier, c'est le cas pour les plantations de cocotiers. L'atténuation de la lumière à l'intérieur d'une feuille résulte de phénomènes complexes liés à sa composition biochimique et à sa structure anatomique alors que l'épiderme détermine surtout la réflectance bidirectionnelle.

Les propriétés optiques des feuilles sont déterminées par les éléments suivants :

- la présence de pigments,
- l'architecture cellulaire de la feuille,
- la physiologie de la feuille.

Les pigments

Les pigments foliaires [Guyot, 1995] sont composés des chlorophylles a et b, des caroténoïdes, des xanthophylles, des plastoquinones et des composés flavoniques.

Une caractéristique importante du végétal est une forte réflectance dans le proche infrarouge due au fait que les pigments foliaires sont transparents aux longueurs d'ondes supérieures à 700 nm.

Dans le visible, le comportement spectral est dû à la composition en pigment parmi lesquels la chlorophylle joue un rôle principal. Les chlorophylles a et b (65% des pigments des plantes supérieures) possèdent deux bandes d'absorption dans le bleu (450 nm) et dans le rouge (660 nm). Cela se traduit par une forte réflectance dans le jaune-vert à une longueur d'onde de 550 nm.

Les seconds pigments par ordre d'importance sont les caroténoïdes et les xanthophylles. Le mieux connu de ces pigments est le β -carotène qui présente un pic d'absorption dans le bleu (450 nm) qui se superpose à celui de la chlorophylle. Ce pic est masqué dans les feuilles saines et vertes. Par contre, ces pigments sont à l'origine du jaunissement des feuilles à l'automne dû à la disparition de la chlorophylle qui cesse alors de masquer la coloration des caroténoïdes.

Structure physiologique de la feuille

La structure de la feuille (voir la figure B.1) joue un rôle important dans le comportement spectral du végétal.

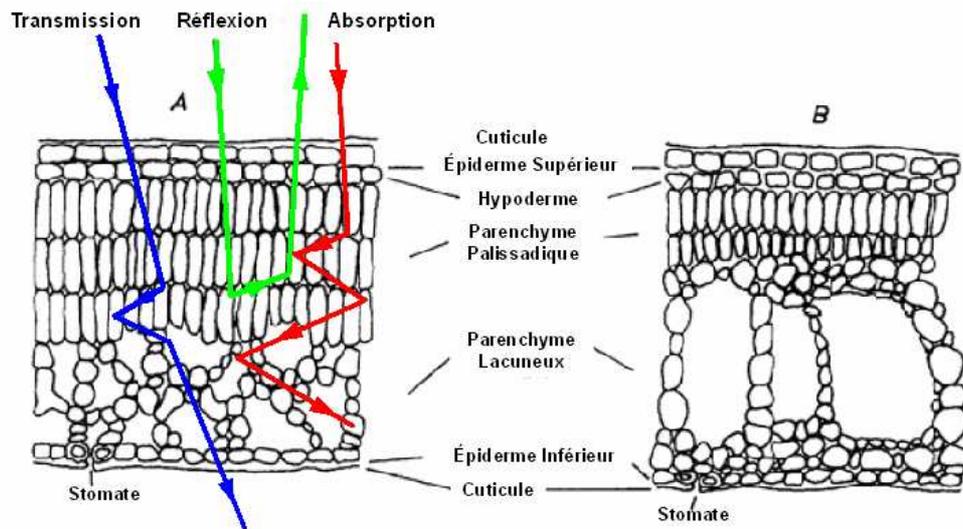


FIG. B.1 – Structure physiologique d'une feuille [Girard and Girard, 1989].

Dans le proche infrarouge (de 700 à 1100 nm), c'est la structure interne du feuillage qui interagit. Le parenchyme lacuneux (ou mésophile) est plus ou moins développé selon les végétaux. Ainsi, un végétal au mésophile bien développé traduit une forte réflectance dans le proche infrarouge. À l'inverse, un végétal dont le mésophile est peu développé réfléchit moins. Il en résulte des différences sensibles dans les courbes de réflectance des différents végétaux. La figure B.2 illustre différentes courbes de réflectance spectrale d'une végétation verte et du sol. La chlorophylle présente dans les feuilles saines absorbe l'énergie solaire à des longueurs d'ondes particulières [Tucker and Sellers, 1986, Swain, 1978]. De plus, l'absorption dans le visible est principalement gouvernée par la composition, la concentration et la distribution des pigments dans les feuilles [Bariou et al., 1985]. Une baisse de réflectance dans le proche infrarouge peut être un signe d'une baisse de vigueur de la plante [Zarco-Tejada et al., 2000, Zarco-Tejada et al., 2004].

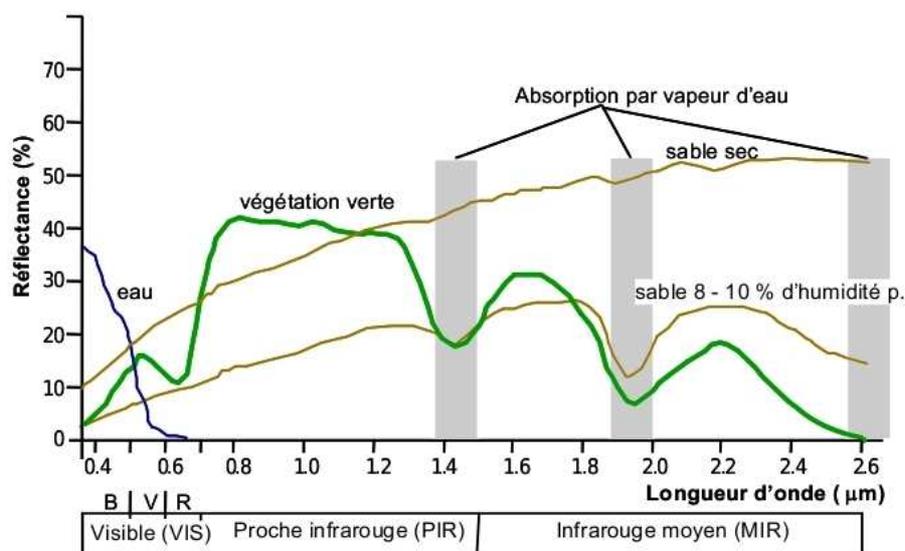


FIG. B.2 – Réponse spectrale des végétaux et du sol.

Comme le montre la figure B.2, la quantité de rayonnement réfléchi est optimale pour la végétation dans le proche infrarouge.

B.2 Les différents indices de végétation

Les différents indices de végétation correspondent à une démarche empirique dans la recherche d'une estimation de paramètres concernant la végétation afin d'identifier des couverts végétaux chlorophylliens, d'évaluer l'indice foliaire, le développement de la canopée, les effets du stress, l'APAR (rayonnement photo synthétiquement actif absorbé), l'évapotranspiration, etc. Ils sont plus ou moins sensibles à la contribution spectrale venant du sol. Ces indices de végétation sont utilisés pour l'estimation de plusieurs paramètres biologiques tels que la biomasse [Tucker, 1979], la concentration de chlorophylle [Buschman and Nagel, 1993], l'indice de surface foliaire [Asrar et al., 1984], les pertes et dommages foliaires [Vogelmann, 1990], l'activité photosynthétique [Sellers, 1985], et bien d'autres encore. La plupart des indices sont basés sur des combinaisons linéaires ou fractionnaires des signaux des différentes bandes radiométriques.

Les indices les plus simple sont les différences VI et les rapports RVI et IP [Knipling, 1970] mais ces types d'indices sont très sensibles aux effets atmosphériques et à la contribution spectrale des sols (pour le RVI).

L'indice de végétation le plus couramment utilisé dans l'agriculture, la gestion forestière, les pâturages et les applications écologiques est l'Indice de Végétation par Différence Normalisée (*Normalized Difference Vegetation Index* ou NDVI) introduit par [Rouse et al., 1974]. Cet indice est basé sur la différence entre l'absorption maximale dans le rouge due aux pigments chlorophylliens et le maximum de la réflexion des radiations dans le proche infrarouge résultant de la structure cellulaire des feuilles. Il est défini de la manière suivante :

$$\text{NDVI} = \frac{\text{PIR} - \text{R}}{\text{PIR} + \text{R}} \quad (\text{B.1})$$

où PIR est la réflectance dans le canal proche infrarouge et R celle dans le canal rouge. La figure B.3 présente l'indice NDVI calculé pour l'ensemble de la planète. Les valeurs inférieures à 0,1 correspondent à des surfaces telles que la roche, le sable ou encore la neige. Les valeurs comprises entre 0,2 et 0,3 indiquent la présence d'arbuste et de régions herbeuses. Les valeurs comprises entre 0,6 et 0,9 permettent de mettre en évidence des régions tempérées et les forêts tropicales.

Il est montré que l'indice NDVI possède plusieurs limitations. Une première concerne la difficulté à distinguer les sombres et/ou humides [Huete, 1987]. Une seconde est la saturation des valeurs de l'indice dans le cas d'une canopée dense et multicouche [Lillesaeter, 1982]. Une troisième est la sensibilité du NDVI aux aérosols présents dans l'atmosphère [Holben, 1986] qui sont caractérisées par l'augmentation apparente de la réflectance dans le rouge due à la diffusion de la lumière du soleil directement sur le capteur et une diminution, à un degré moindre, de la réflectance dans la proche infra-rouge par absorption de la lumière du soleil.

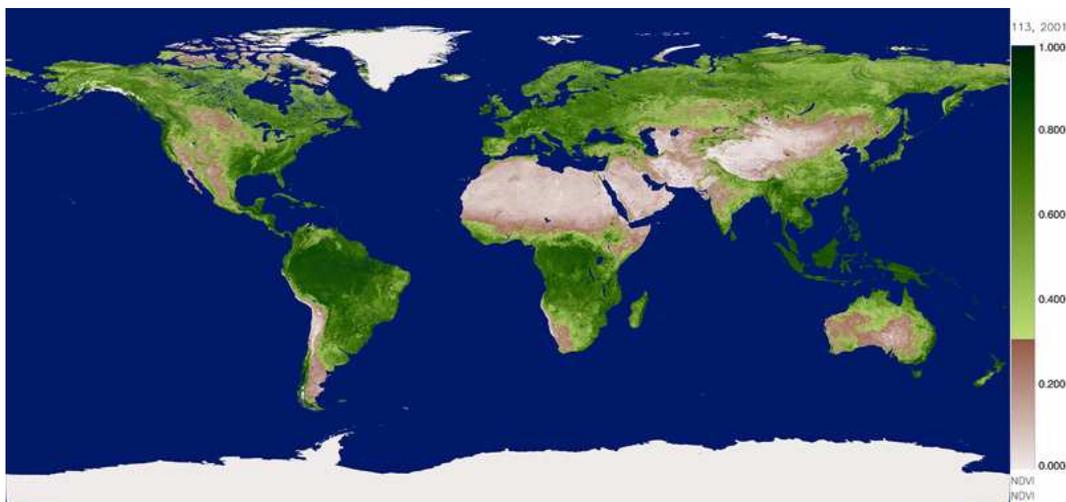


FIG. B.3 – Composition colorée du NDVI calculé sur l'ensemble de la Terre.

Ceci a motivé le développement de nouveaux indices pour minimiser l'influence des sols. Citons le

SAVI (*Soil Adjusted Vegetation Index*) introduit par [Huete, 1988], SAVI transformé (*Transformed SAVI*) [Baret et al.,], le SAVI2 (*Soil Adjusted Ratio Vegetation Index*) [Major et al., 1990] ou encore le MSAVI (*Modified SAVI*) [Qi et al., 1994]. D'autres indices de végétation ont été proposés pour corriger directement l'effet des aérosols dans la bande rouge en incorporant la bande bleue. Ces indices sont l'ARVI (*Atmospheric Resistant Vegetation Index*) et le SARVI (*Soil-Adjusted and Atmospheric Resistant Vegetation Index*), tous deux introduits par [Kaufman and Tanre, 1992]. Le TVI (*Transformed Vegetation Index*) développé par [Deering et al., 1975] essaie d'éliminer les valeurs négatives du NDVI et stabilise la variance.

En exploitant la philosophie de la droite des sols nus (pente et origine), [Richardson and Wiegand, 1977] ont proposé le PVI (*Perpendicular Vegetation Index*). La transformation du chapeau à corne (*Tasseled Cap*) [Kauth and Thomas, 1976] est une transformation orthogonale des 4 canaux pour réduire la sensibilité à la contribution spectrale des sols, sans, malgré tout, pouvoir l'éliminer complètement. Une liste non exhaustive d'indices de végétation les plus utilisés est présentée dans le tableau B.1

Bandes	Sigle	
Bande rouge	R	acquisition dans le spectre visible du rouge
Bande vert	V	acquisition dans le spectre visible du vert
Bande bleu	B	acquisition dans le spectre visible du bleu
Proche Infrarouge	PIR	acquisition dans le proche infrarouge
Infrarouge	IR	acquisition dans l'infrarouge
Dénomination	Formule	
Différence	VI	$R - \text{PIR}$
Rapport	RVI	$\frac{\text{PIR}}{R}$
	IP	$\frac{V}{R}$
Indice de végétation normalisé	NDVI	$\frac{\text{PIR} - R}{\text{PIR} + R}$
Indice de végétation transformé	TVI	$\sqrt{\text{NDVI}} + 0,5$
Indice de végétation perpendiculaire	PVI	$a_1 \text{PIR} - a_2 R + \alpha$
Chapeau à corne (<i>"Tasseled Cap"</i>)	Formule générale $a_1 V + a_2 R + a_3 \text{PIR} + a_4 \text{PIR}$	
Indice de verdeur issu du précédent	GVI	$b_1 V - b_2 R + b_3 \text{PIR} + b_4 \text{PIR}$
Indice de végétation ajusté au sol	SAVI	$\frac{(1+L)(\text{PIR} - R)}{(\text{PIR} + R + L)}$ $L = 0,5$ pour diminuer l'effet du sol Si $L = 0$ alors SAVI = NDVI
Indice de végétation normalisé corrigé des effets atmosphérique	ARVI	$(\text{PIR} - \text{RB}) / (\text{PIR} + \text{RB})$ $\text{RB} = R - \gamma(B - R)$ B R réflectance dans le bleu et le rouge, γ une fonction du type aérosols
Indice de végétation amélioré	EVI	$2.5 \frac{\text{PIR} - R}{\text{PIR} + 6R - 7.5B + 1}$

TAB. B.1 – Indices de végétation les plus couramment utilisés adapté de [Girard and Girard, 1989].

Génération des masques de végétation

Dans cette annexe, il est présenté les valeurs utilisées lors du processus de segmentation des zones de végétation haute ainsi que des planches illustrant la méthode utilisée pour certains *motu*.

Tableau récapitulatif des valeurs utilisées

Le tableau C.1 présente, pour chaque *motu*, les différentes valeurs utilisées pour la segmentation et la classification de la végétation haute.

Étapes Motu	ROI	Sieve seuil min	Clump fenêtre	Masque Végétal	Seuillage	Morpho Math	Ajustement
1	3 ROIs : - Lagon - Végétation - Sable-Corail	10	3x3	Classe Vég. min pop. : 10000	DN=40 Min pop. : 100000	Fermeture : 5x5	Remasquer avec masque végétal
2	4 ROIs : - Lagon - Végétation - Eau peu pro- fonde	10	5x5	Classe Vég. min pop. : 10000	DN=60 Min pop. : 1000	Fermeture : 11x11 Ouverture : 7x7	
3	3 ROIs : - Lagon - Végétation - Eau peu pro- fonde	10	5x5	Classe Vég. min pop. : 10000	DN=60 Min pop. : 10000	Fermeture : 11x11 Ouverture : 11x11	Remasquer avec masque végétal fermeture 7x7
4	4 ROIs : - Lagon - Végétation - Eau peu pro- fonde - Sable - Corail	10	5x5	Classe Vég. min pop. : 10000	DN=50 Min pop. : 10000	Fermeture : 9x9 Ouverture : 9x9	Resegmenter résultat min.DN=1 et min pop.=10000
5	4 ROIs : - Lagon - Végétation - Eau peu pro- fonde - Sable - Corail	10	3x3	Classe Vég. min pop. : 10000	DN=40 Min pop. : 10000	Fermeture : 11x11 Ouverture : 11x11	
6	5 ROIs : - Lagon - Végétation - Eau peu pro- fonde - Sable - Corail	10	5x5	Classe Vég. min pop. : 10000	DN=160 Min pop. : 1000	Fermeture : 11x11 Ouverture : 7x7	
7	NON TRAITÉ						
8	NON TRAITÉ						
9	NON TRAITÉ						
10	4 ROIs : - Lagon - Végétation - Eau peu pro- fonde - Sable	10	3x3	Classe Vég. min pop. : 10000	DN=200 Min pop. : 1000	Fermeture : 9x9	
11	5 ROIs : - Lagon - Végétation - Eau profonde - Eau peu pro- fonde - Sable	10	7x7	Classe Vég. min pop. : 300000	DN=180 Min pop. : 1000	Fermeture : 13x13 Ouverture : 7x7	
12	X ROIs	10	5x5	Classe Vég. min pop. : 10000 re-segmenter 2 = DN = 3	DN=140 Min pop. : 1000	Fermeture : 13x13 Ouverture : 5x5	
13	X ROIs	10	5x5	Classe Vég.	DN=140		
14	X ROIs	10	5x5	Classe Vég.	DN=140		

15	2 ROIs : - Végétation - Sable	2	3x3	Classe Vég. min pop. : 10000	Pas de segmentation de texture nécessaire	Fermeture : 3x3	
16	5 ROIs : - Lagon - Végétation - Eau profonde - Eau peu profonde - Sable	10	3x3	Classe Vég. min pop. : 400000	DN=140		
17	3 ROIs : - Lagon - Végétation - Sable	2	3x3	Classe Vég. min pop. : 1000	DN=140	Fermeture : 5x5 Ouverture : 3x3	
18	3 ROIs : - Lagon - Végétation - Sable	2	3x3	Classe Vég. min pop. : 1000	DN=140	Fermeture : 5x5 Ouverture : 3x3	
19	4 ROIs : - Lagon - Végétation - Sable - Autre	2	3x3	Classe Vég. min pop. : 2000	Pas de segmentation de texture nécessaire	Fermeture : 5x5 Ouverture : 7x7	
20	4 ROIs : - Lagon - Végétation - Sable - Autre	2	3x3	Classe Vég. min pop. : 1000		Fermeture : 3x3 Ouverture : 5x5	
21	X ROIs	2	3x3	Classe Vég. min pop. : 1000		Fermeture : 3x3 Ouverture : 5x5	
22	4 ROIs : - Lagon - Végétation - Sable-Corail	2	3x3	Classe Vég. min pop. : 1000			
23	X ROIs	2	3x3	Classe Vég. min pop. : 1000		Fermeture : 3x3 Ouverture : 5x5	
23	4 ROIs : - Lagon - Végétation - Sable-Corail - Eau peu profonde	2	3x3	Classe Vég. min pop. : 1000	DN=140 Min pop. :100	Fermeture : 9x9 Ouverture : 3x3	

TAB. C.1: Récapitulatif des étapes de génération des masques pour chacun des motus traités.

Le motu 01

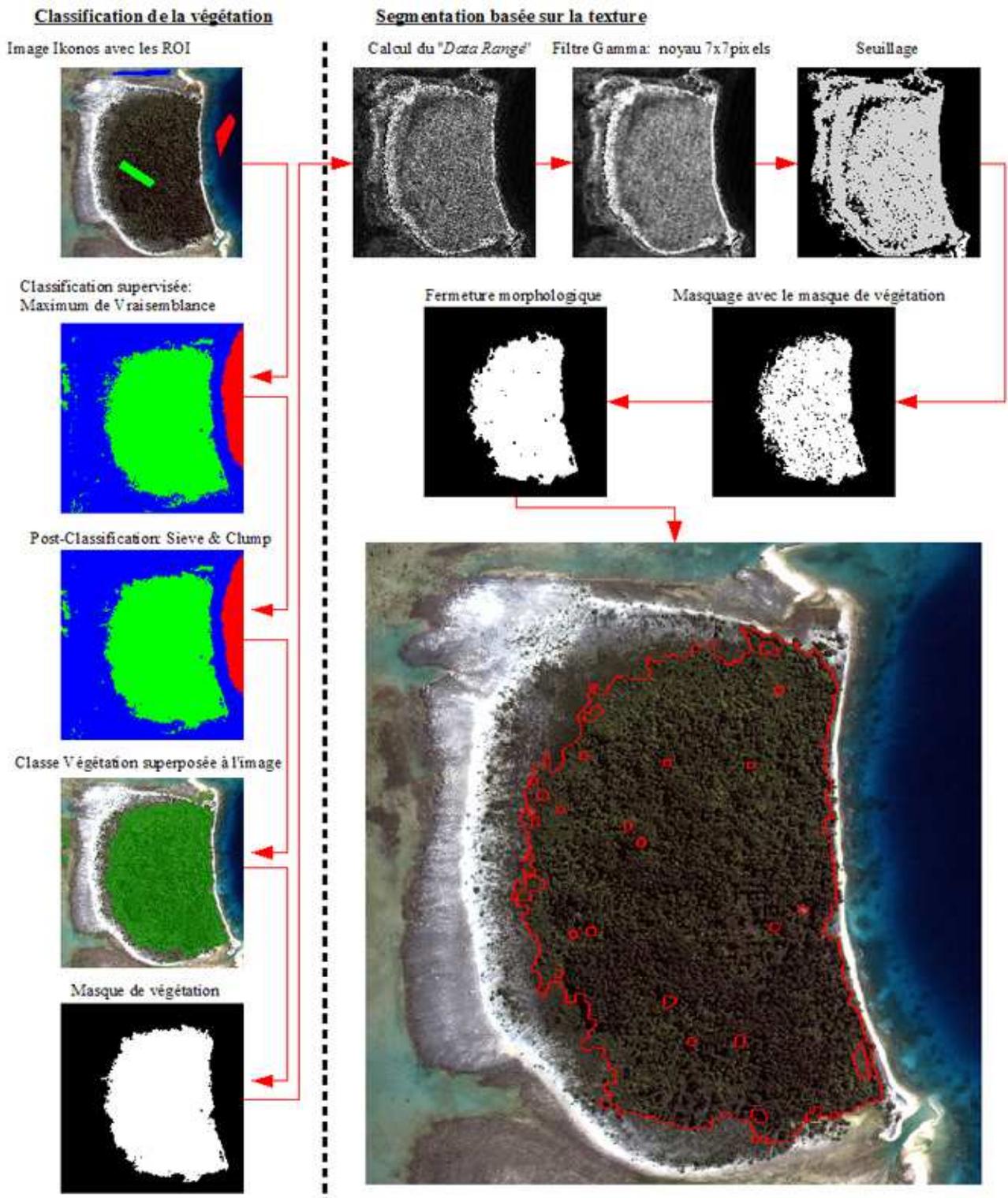


FIG. C.1 – Génération du masque de la végétation haute.

Le motu 02

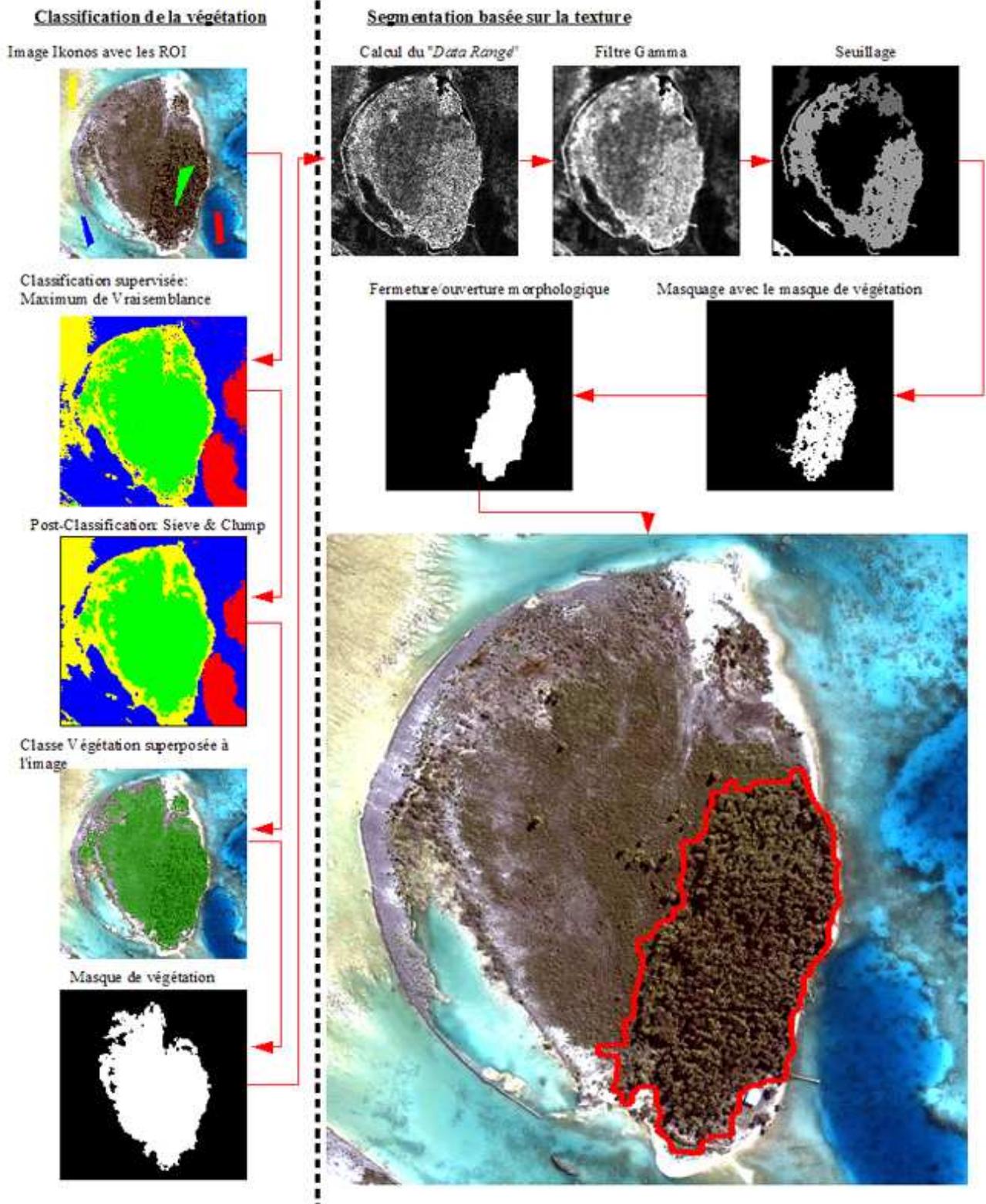


FIG. C.2 – Génération du masque de la végétation haute.

Le motu 03

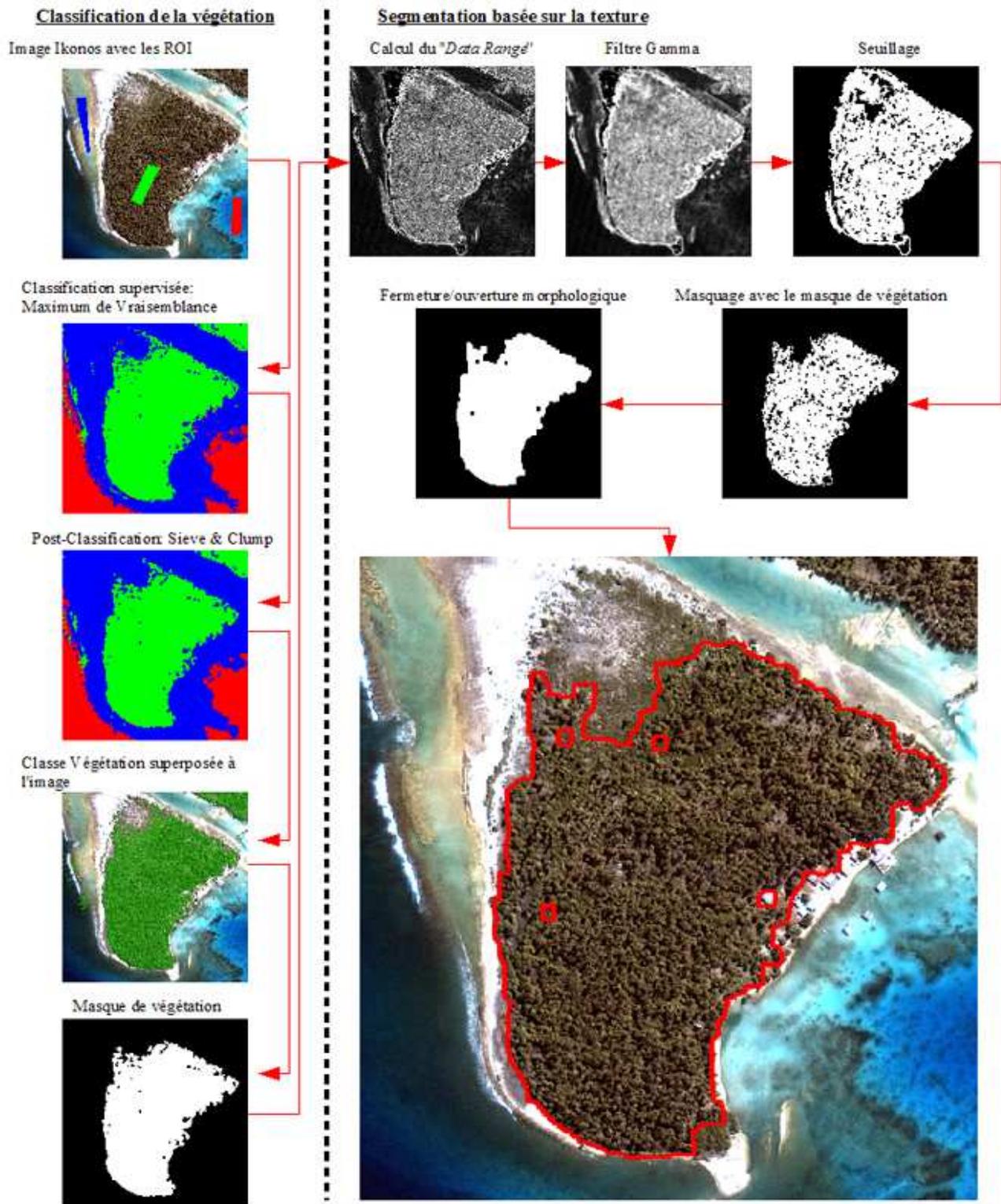


FIG. C.3 – Génération du masque de la végétation haute.

Le motu 04

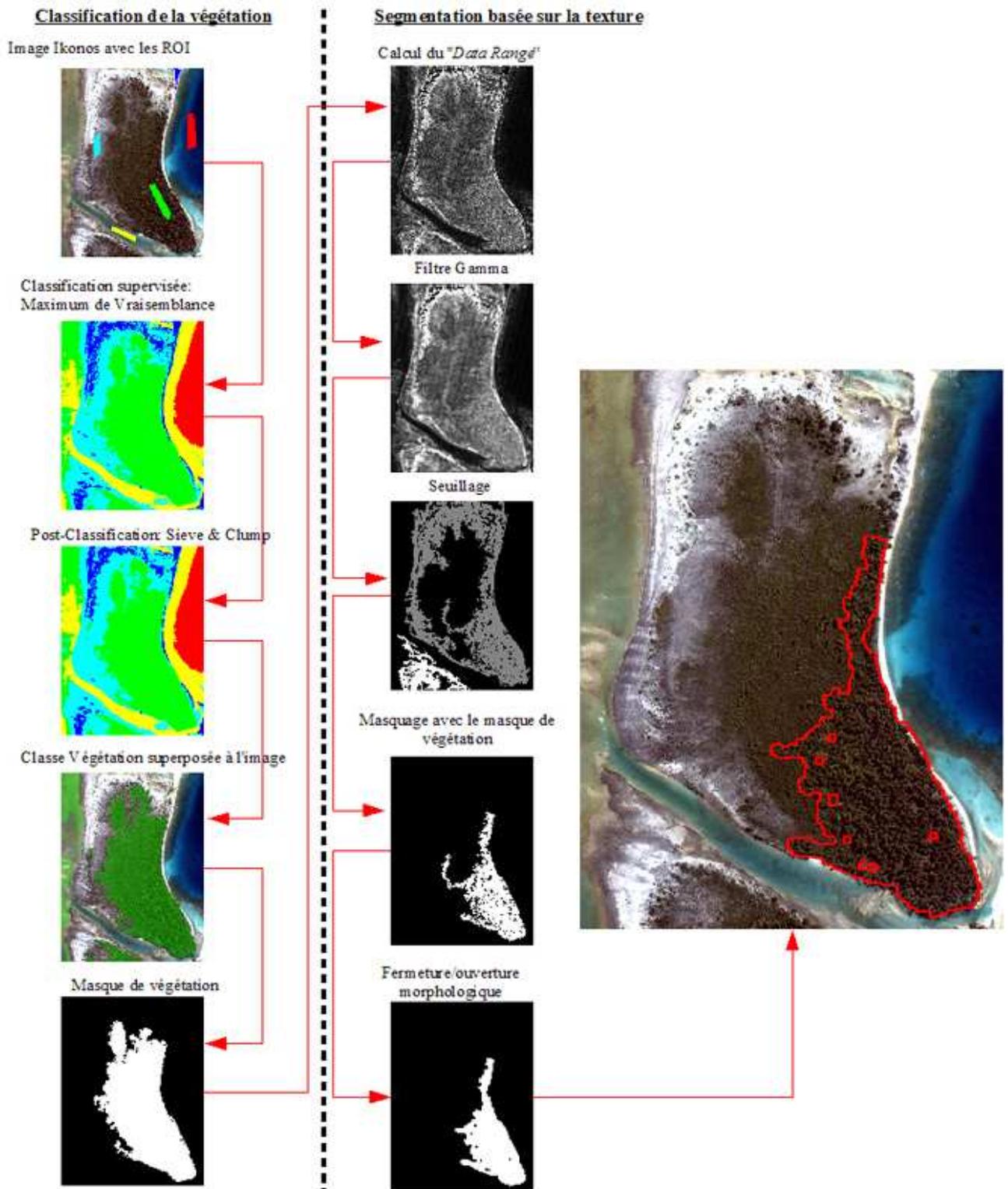


FIG. C.4 – Génération du masque de la végétation haute.

Le motu 05

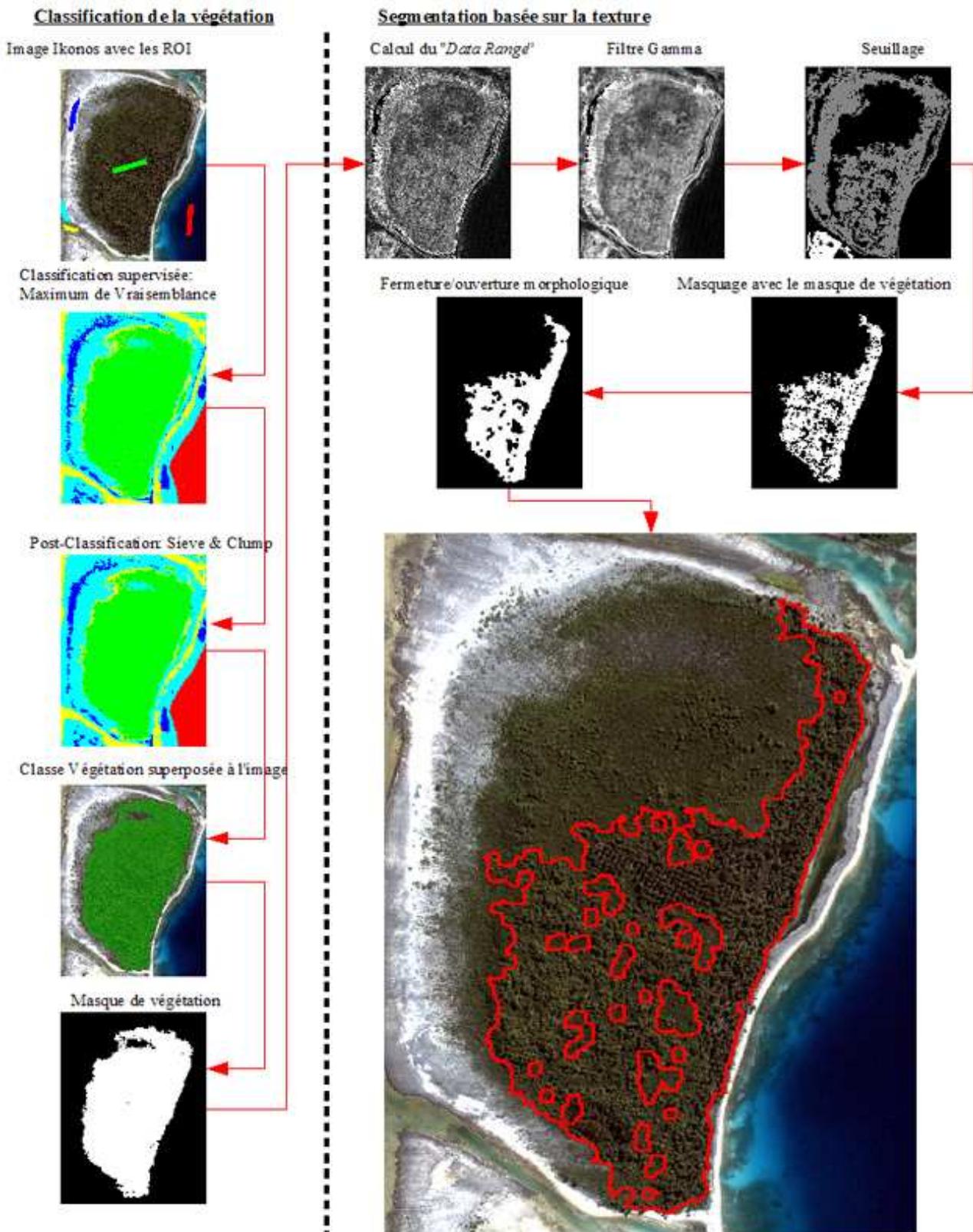


FIG. C.5 – Génération du masque de la végétation haute.

Le motu 06

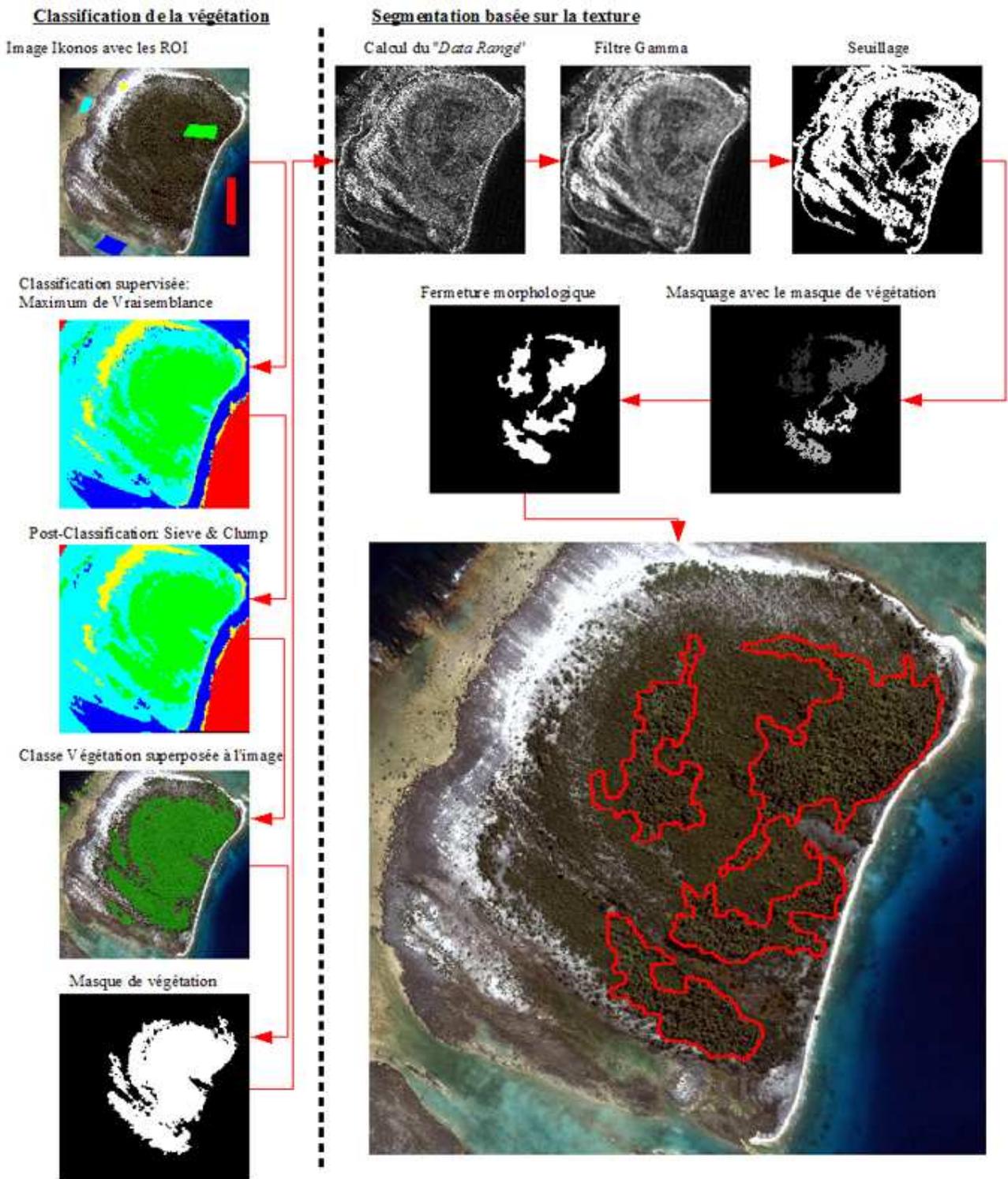


FIG. C.6 – Génération du masque de la végétation haute.

Le motu 10

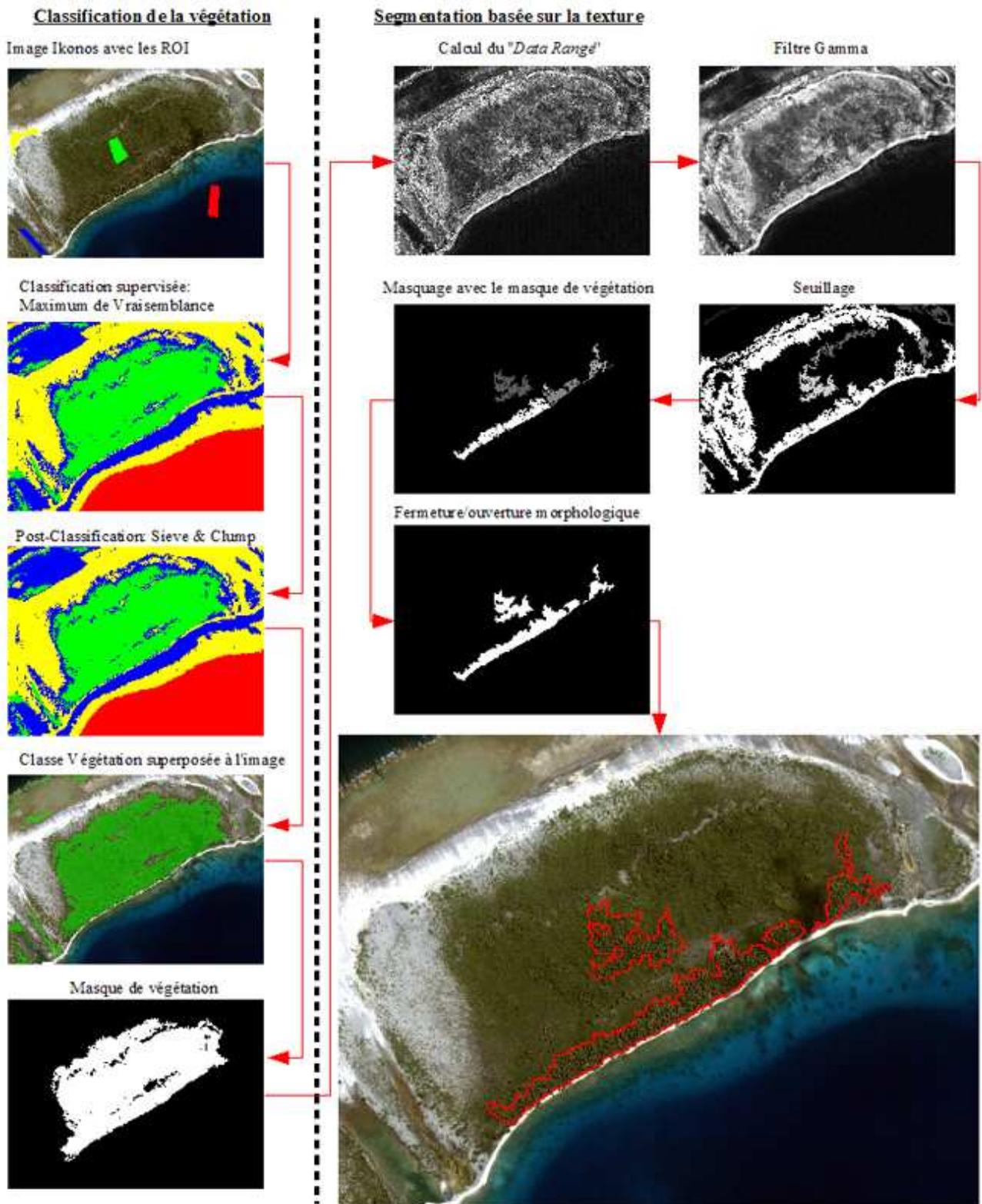


FIG. C.7 – Génération du masque de la végétation haute.

Le motu 11

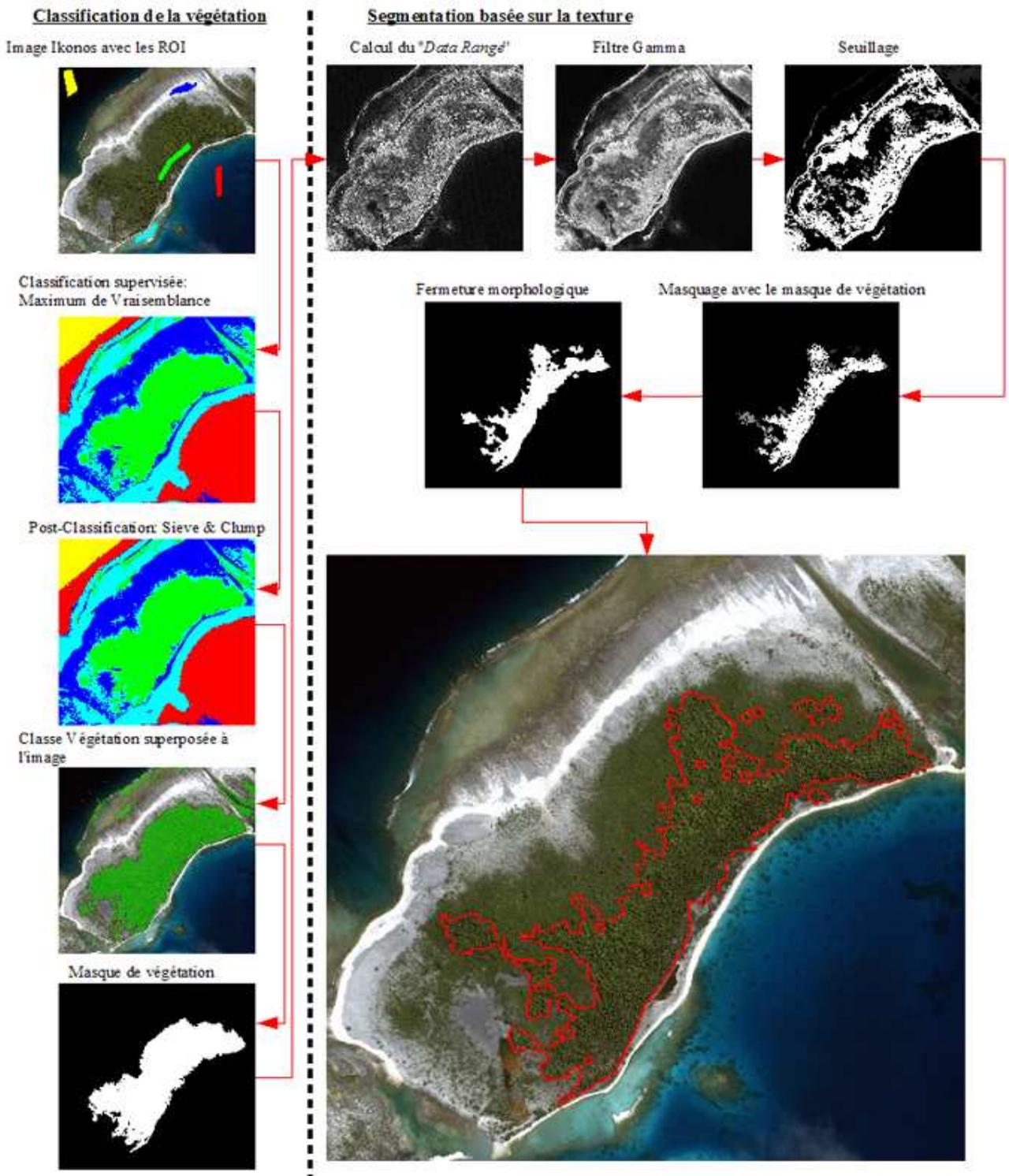


FIG. C.8 – Génération du masque de la végétation haute.

Le motu 12

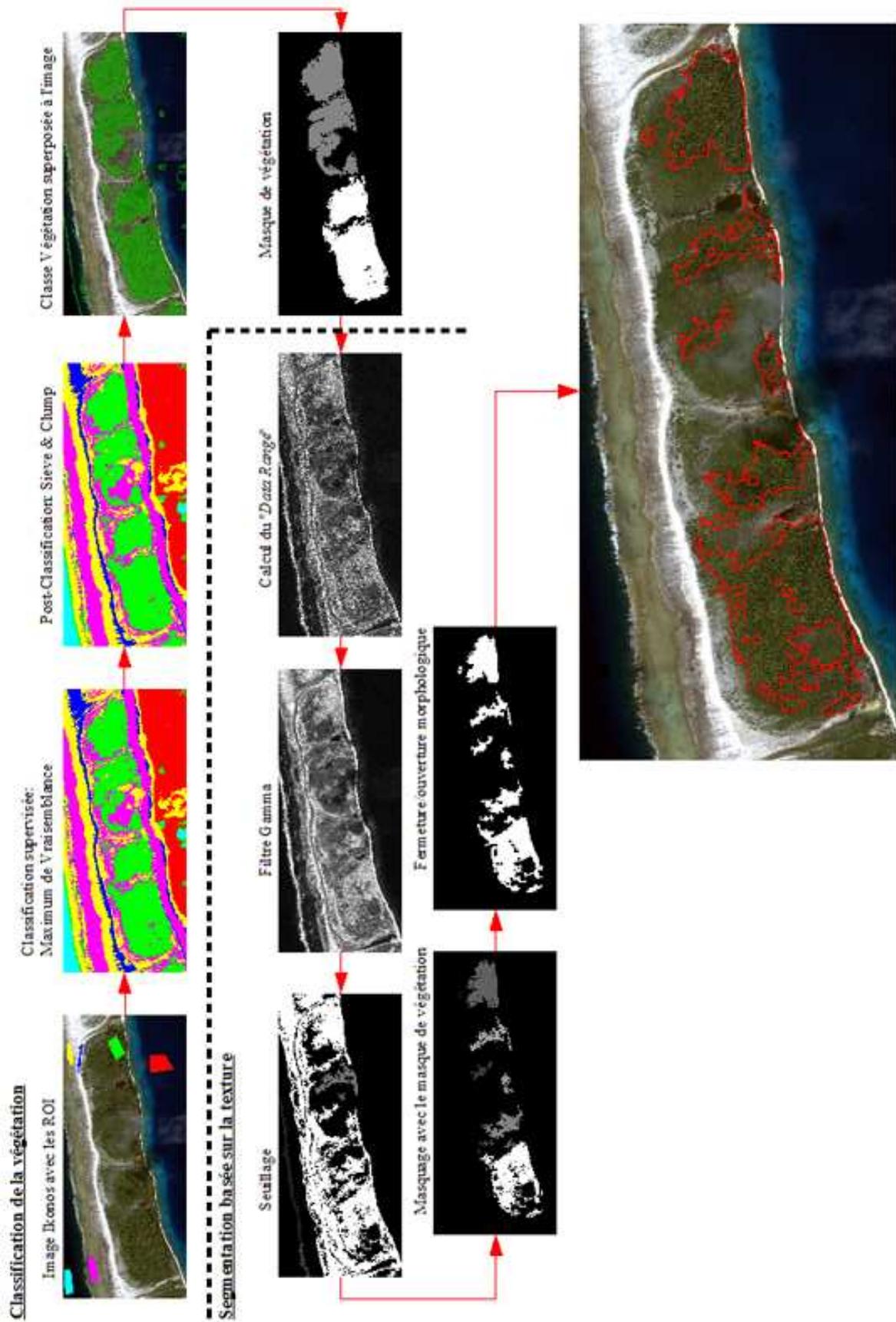


FIG. C.9 – Génération du masque de la végétation haute.

Le motu 14

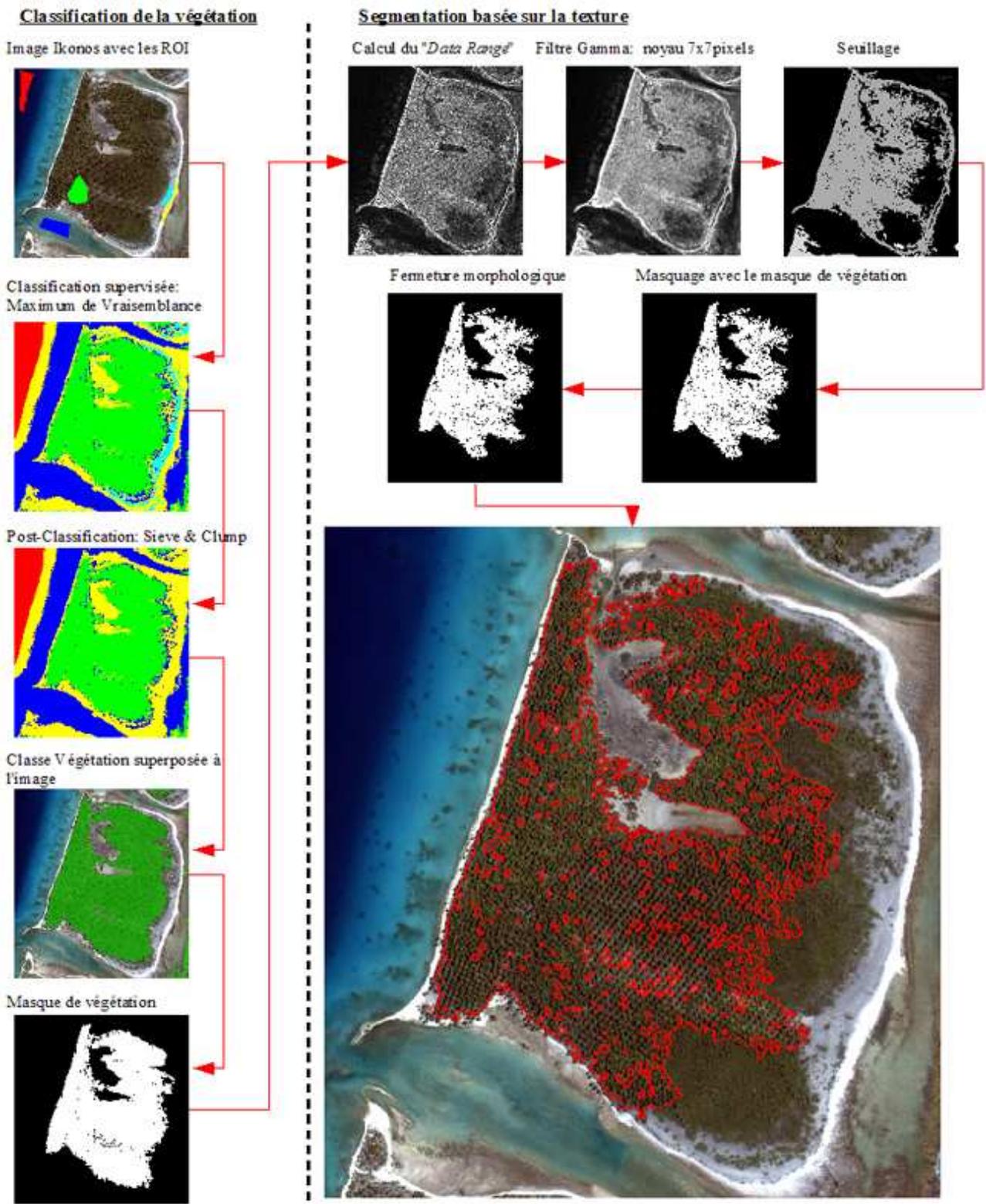


FIG. C.10 – Génération du masque de la végétation haute.

Le motu 15

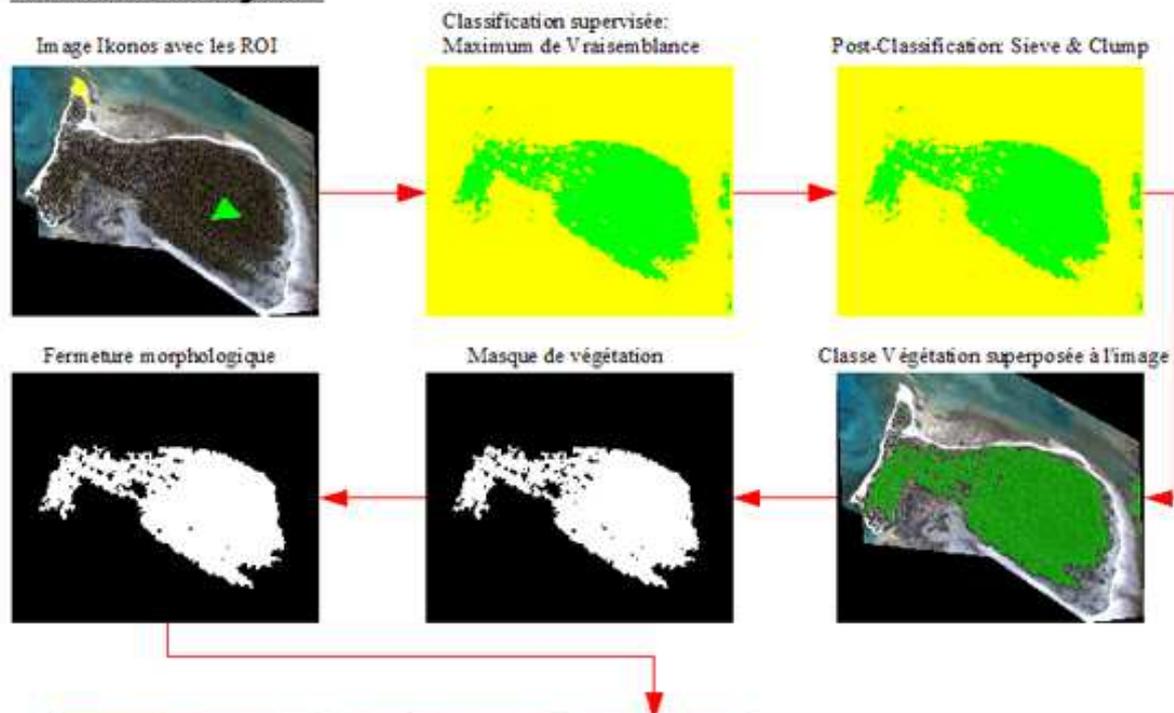
Classification de la végétation

FIG. C.11 – Génération du masque de la végétation haute.

Cartes de densité des secteurs test

Cette annexe présente les différentes cartes de densité locale générée à partir des résultats de la segmentation individuelle des arbres sur les secteurs test du *Motu*#16.

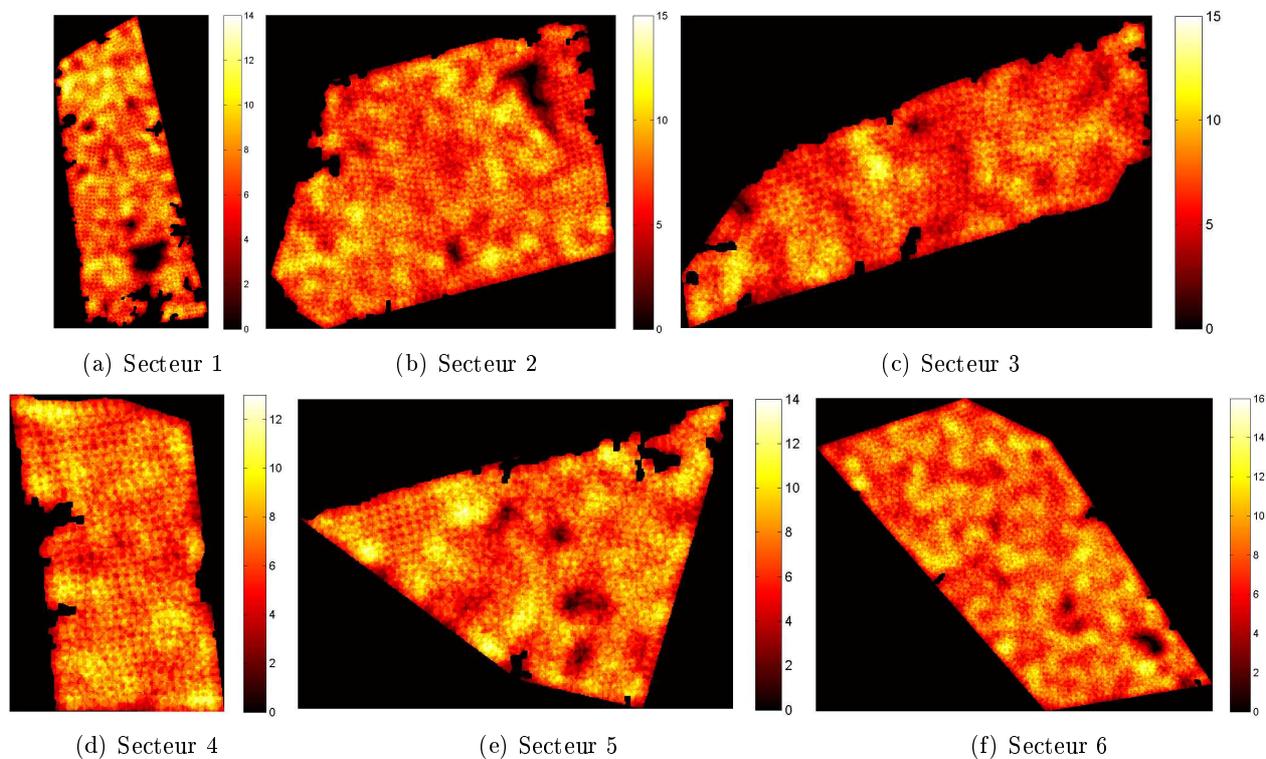


FIG. D.1 – Cartes de densité des secteurs de test du motu 16.

Méthode utilisée pour l'interprétation visuelle

Pour réaliser le travail de vérité par photo interprétation humaine, nous avons utilisé logiciel libre GIMP (voir la figure E.1).

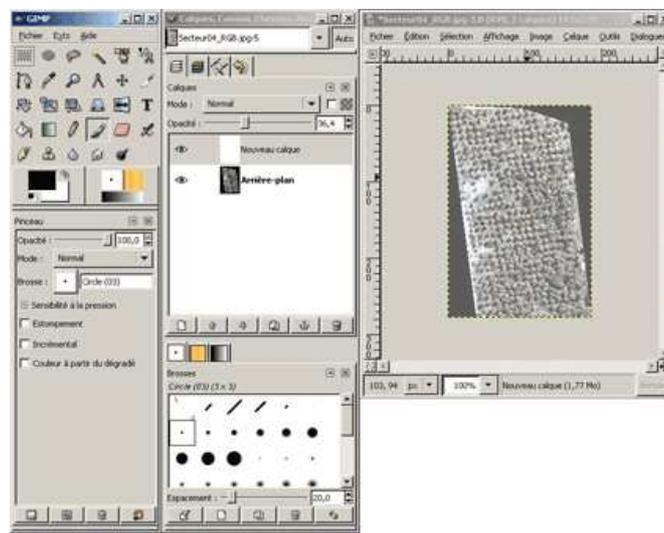


FIG. E.1 – Interface de GIMP.

Pour chaque secteur de test du motu 16, nous avons créé un calque avec un fond blanc sur lequel nous avons placé des points noirs au centre de chaque objet que nous avons interprété comme étant un cocotier. En modifiant l'opacité du calque, il est possible de voir et de placer ces points noirs tout en contrôlant qu'ils ne se touchent pas. La figure E.2 montre un secteur et son calque correspondant visible après modification de l'opacité.

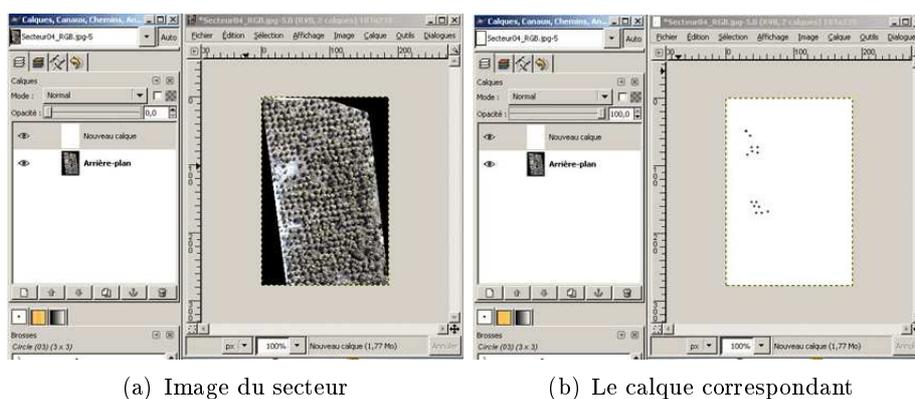


FIG. E.2 – Exemple d'un secteur et du calque en jouant avec l'opacité de ce dernier.

La forme d'un cocotier est très caractéristique, elle est circulaire. La figure E.3 montre un exemple de cocotiers présents dans l'image.



FIG. E.3 – Exemple d'arbres visibles sur les images.

La principale difficulté de ce travail est qu'à force d'observer les mêmes motifs d'une image (texture relativement homogène), l'oeil humain se fatigue et nous avons alors du mal à interpréter correctement ce que nous voyons. Il faut respecter un long temps de repos avant l'examen d'une nouvelle image. C'est donc un travail long. Une fois les calques réalisés, un petit programme dénombre automatiquement les points noirs.

Annexe **F**

Correspondance des résultats avec les mesures sur le terrain

Dans cette annexe, il est présenté les différents tableaux des relevés effectués sur le terrain ainsi que la mise en correspondance avec les résultats de la méthode.

Secteur02 :			Terrain			Densité			H3 DE (S1+S2+S3)			NDA (S2+S3)			H1 NDE			TDP (S1)			NDA (S3)			H2 NDE			TDP (S1+S2)		
IdPlacette	S1	S2	S3	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.		
1	4	1	0	8	2	4,93	160%	40%	99%	20%	40%	-80%	-19%	200%	50%	123%	0%	60%	60%	1%	160%	40%	99%	160%	40%	99%			
2	8	3	2	8	4	6,11	62%	31%	47%	38%	31%	0%	15%	100%	50%	76%	15%	54%	54%	38%	73%	36%	56%	73%	36%	56%			
3	6	4	0	10	4	7,19	100%	40%	72%	40%	20%	-40%	-12%	167%	67%	120%	0%	60%	60%	28%	100%	40%	72%	100%	40%	72%			
4	8	2	2	9	5	6,97	75%	42%	58%	33%	25%	-8%	9%	113%	63%	87%	17%	42%	42%	25%	90%	50%	70%	90%	50%	70%			
5	8	1	1	8	3	5,32	80%	30%	53%	20%	50%	0%	27%	100%	38%	67%	10%	60%	60%	37%	89%	33%	59%	89%	33%	59%			
6	6	3	0	9	3	6,36	100%	33%	71%	33%	33%	-33%	-4%	150%	50%	106%	0%	67%	67%	29%	100%	33%	71%	100%	33%	71%			
7	4	6	6	9	5	6,50	56%	31%	41%	75%	-6%	-31%	-16%	225%	125%	163%	38%	31%	31%	22%	90%	50%	65%	90%	50%	65%			
8	5	2	1	11	6	7,79	138%	75%	97%	38%	-13%	-75%	-35%	220%	120%	156%	13%	13%	13%	-10%	157%	86%	111%	157%	86%	111%			
9	9	4	1	8	5	6,36	57%	36%	45%	36%	29%	7%	19%	89%	56%	71%	7%	57%	57%	47%	62%	38%	49%	62%	38%	49%			
10	6	3	2	11	6	8,36	100%	55%	76%	45%	0%	-45%	-21%	183%	100%	139%	18%	27%	27%	6%	122%	67%	93%	122%	67%	93%			

TAB. F.1 – Récapitulatif des mises en correspondance des relevés terrain avec les cartes de densités générées à partir des résultats de détection de la méthode pour le secteur de test 2.

Secteur03 :			Terrain			Densité			H3 DE (S1+S2+S3)			NDA (S2+S3)			H1 NDE			TDP (S1)			NDA (S3)			H2 NDE			TDP (S1+S2)		
IdPlacette	S1	S2	S3	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.		
1	7	2	0	7	3	5,28	78%	33%	59%	22%	44%	0%	19%	100%	43%	75%	0%	67%	67%	41%	78%	33%	59%	78%	33%	59%			
2	5	1	0	6	2	4,14	100%	33%	69%	17%	50%	-17%	14%	120%	40%	83%	0%	67%	67%	31%	100%	33%	69%	100%	33%	69%			
3	8	3	1	10	5	7,43	83%	42%	62%	33%	25%	-17%	5%	125%	63%	93%	8%	50%	50%	30%	91%	45%	68%	91%	45%	68%			
4	7	5	0	9	4	6,92	75%	33%	58%	42%	25%	-17%	1%	129%	57%	99%	0%	67%	67%	42%	75%	33%	58%	75%	33%	58%			
5	6	4	1	10	4	7,17	91%	36%	65%	45%	18%	-36%	-11%	167%	67%	119%	9%	55%	55%	26%	100%	40%	72%	100%	40%	72%			
6	5	1	0	10	6	7,90	167%	100%	132%	17%	-17%	-83%	-48%	200%	120%	158%	0%	0%	0%	-32%	167%	100%	132%	167%	100%	132%			

TAB. F.2 – Récapitulatif des mises en correspondance des relevés terrain avec les cartes de densités générées à partir des résultats de détection de la méthode pour le secteur de test 3.

Secteur04 :			Terrain			Densité			H3 DE (S1+S2+S3)			NDA (S2+S3)			H1 NDE			TDP (S1)			NDA (S3)			H2 NDE			TDP (S1+S2)		
IdPlacette	S1	S2	S3	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.		
1	7	1	0	9	2	5,49	113%	25%	69%	13%	63%	-25%	19%	129%	29%	78%	0%	75%	75%	31%	113%	25%	69%	113%	25%	69%			
2	6	1	0	7	4	5,60	100%	57%	80%	14%	29%	-14%	6%	117%	67%	93%	0%	43%	43%	20%	100%	57%	80%	100%	57%	80%			
3	5	0	0	9	3	6,38	180%	60%	128%	0%	40%	-80%	-28%	180%	60%	128%	0%	40%	40%	-28%	180%	60%	128%	180%	60%	128%			
4	6	2	0	8	4	5,67	100%	50%	71%	25%	25%	-25%	4%	133%	67%	95%	0%	50%	50%	29%	100%	50%	71%	100%	50%	71%			
5	7	4	0	7	2	4,94	64%	18%	45%	36%	45%	0%	19%	100%	29%	71%	0%	82%	82%	55%	64%	18%	45%	64%	18%	45%			
6	5	0	0	9	4	6,23	180%	80%	125%	0%	20%	-80%	-25%	180%	80%	125%	0%	20%	20%	-25%	180%	80%	125%	180%	80%	125%			
7	3	4	3	7	2	4,31	70%	20%	43%	70%	10%	-40%	-13%	233%	67%	144%	30%	50%	50%	27%	100%	20%	62%	100%	20%	62%			
8	7	1	0	7	4	4,05	88%	50%	51%	13%	38%	0%	37%	100%	57%	58%	0%	50%	50%	49%	88%	50%	51%	88%	50%	51%			
9	5	2	0	6	2	4,07	86%	29%	58%	29%	43%	-14%	13%	120%	40%	81%	0%	71%	71%	42%	86%	29%	58%	86%	29%	58%			
10	6	2	0	10	5	7,66	125%	63%	96%	25%	13%	-50%	-21%	167%	83%	128%	0%	38%	38%	4%	125%	63%	96%	125%	63%	96%			
11	5	1	0	8	4	5,78	133%	67%	96%	17%	17%	-50%	-13%	160%	80%	116%	0%	33%	33%	4%	133%	67%	96%	133%	67%	96%			
12	6	4	0	9	3	5,70	90%	30%	57%	40%	30%	-30%	3%	150%	50%	95%	0%	70%	70%	43%	90%	30%	57%	90%	30%	57%			
13	6	2	0	8	3	6,04	100%	38%	76%	25%	38%	-25%	-1%	133%	50%	101%	0%	63%	63%	24%	100%	38%	76%	100%	38%	76%			
14	4	4	0	8	4	6,39	100%	50%	80%	50%	0%	-50%	-30%	200%	100%	160%	0%	50%	50%	20%	100%	50%	80%	100%	50%	80%			
15	7	1	0	7	3	5,11	88%	38%	64%	13%	50%	0%	24%	100%	43%	73%	0%	63%	63%	36%	88%	38%	64%	88%	38%	64%			

TAB. F.3 – Récapitulatif des mises en correspondance des relevés terrain avec les cartes de densités générées à partir des résultats de détection de la méthode pour le secteur de test 4.

Secteur05 :				Terrain			Densité			H3 DE (S1+S2+S3)			NDA (S2+S3)	H1 NDE			TDP (S1)	NDA (S3)	H2 NDE			TDP (S1+S2)		
IdPlacette	S1	S2	S3	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.
1	11	4	8	9	4	6,42	39%	17%	28%	52%	30%	9%	20%	82%	36%	58%	35%	48%	48%	37%	60%	27%	43%	
2	14	4	0	12	5	7,99	67%	28%	44%	22%	50%	11%	33%	86%	36%	57%	0%	72%	72%	56%	67%	28%	44%	
3	8	3	2	9	4	6,65	69%	31%	51%	38%	31%	-8%	10%	113%	50%	83%	15%	54%	54%	33%	82%	36%	60%	
4	14	0	1	13	6	9,40	87%	40%	63%	7%	53%	7%	31%	93%	43%	67%	7%	53%	53%	31%	93%	43%	67%	
5	6	2	6	8	3	5,01	57%	21%	36%	57%	21%	-14%	7%	133%	50%	83%	43%	36%	36%	21%	100%	38%	63%	
6	8	3	0	9	4	6,17	82%	36%	56%	27%	36%	-9%	17%	113%	50%	77%	0%	64%	64%	44%	82%	36%	56%	
7	5	1	2	9	2	5,09	113%	25%	64%	38%	38%	-50%	-1%	180%	40%	102%	25%	50%	50%	11%	150%	33%	85%	

TAB. F.4 – Récapitulatif des mises en correspondance des relevés terrain avec les cartes de densités générées à partir des résultats de détection de la méthode pour le secteur de test 5.

Secteur05 :				Terrain			Densité			H3 DE (S1+S2+S3)			NDA (S2+S3)	H1 NDE			TDP (S1)	NDA (S3)	H2 NDE			TDP (S1+S2)		
IdPlacette	S1	S2	S3	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.	max	min	moy.
1	7	4	2	11	6	8,24	85%	46%	63%	46%	8%	-31%	-10%	157%	86%	118%	15%	38%	38%	21%	100%	55%	75%	
2	8	7	2	11	6	8,31	65%	35%	49%	53%	12%	-18%	-2%	138%	75%	104%	12%	53%	53%	39%	73%	40%	55%	
3	10	4	0	11	6	8,43	79%	43%	60%	29%	29%	-7%	11%	110%	60%	84%	0%	57%	57%	40%	79%	43%	60%	
4	5	4	0	8	3	5,69	89%	33%	63%	44%	22%	-33%	-8%	160%	60%	114%	0%	67%	67%	37%	89%	33%	63%	
5	6	3	0	10	4	6,65	111%	44%	74%	33%	22%	-44%	-7%	167%	67%	111%	0%	56%	56%	26%	111%	44%	74%	
6	8	7	0	10	3	7,16	67%	20%	48%	47%	33%	-13%	6%	125%	38%	90%	0%	80%	80%	52%	67%	20%	48%	
7	9	5	2	11	3	7,40	69%	19%	46%	44%	38%	-13%	10%	122%	33%	82%	13%	69%	69%	41%	79%	21%	53%	
8	6	4	0	13	6	9,70	130%	60%	97%	40%	0%	-70%	-37%	217%	100%	162%	0%	40%	40%	3%	130%	60%	97%	
9	9	5	2	11	5	8,06	69%	31%	50%	44%	25%	-13%	6%	122%	56%	90%	13%	56%	56%	37%	79%	36%	58%	
10	7	3	2	10	4	6,84	83%	33%	57%	42%	25%	-25%	1%	143%	57%	98%	17%	50%	50%	26%	100%	40%	68%	
11	5	4	1	9	4	6,09	90%	40%	61%	50%	10%	-40%	-11%	180%	80%	122%	10%	50%	50%	29%	100%	44%	68%	
12	8	4	1	10	5	7,28	77%	38%	56%	38%	23%	-15%	6%	125%	63%	91%	8%	54%	54%	36%	83%	42%	61%	
13	5	1	3	11	6	8,11	122%	67%	90%	44%	-11%	-67%	-35%	220%	120%	162%	33%	0%	0%	-23%	183%	100%	135%	

TAB. F.5 – Récapitulatif des mises en correspondance des relevés terrain avec les cartes de densités générées à partir des résultats de détection de la méthode pour le secteur de test 6.

Cas des *motu* 13 et *motu* 16

G.1 *motu* 13

Étant donné l'étendue du *motu* et la présence d'eau en grande majorité et dans un souci de réduction des temps de calcul, nous avons partitionné le *motu* en 4 parties et chacune de ces parties en imagerie.

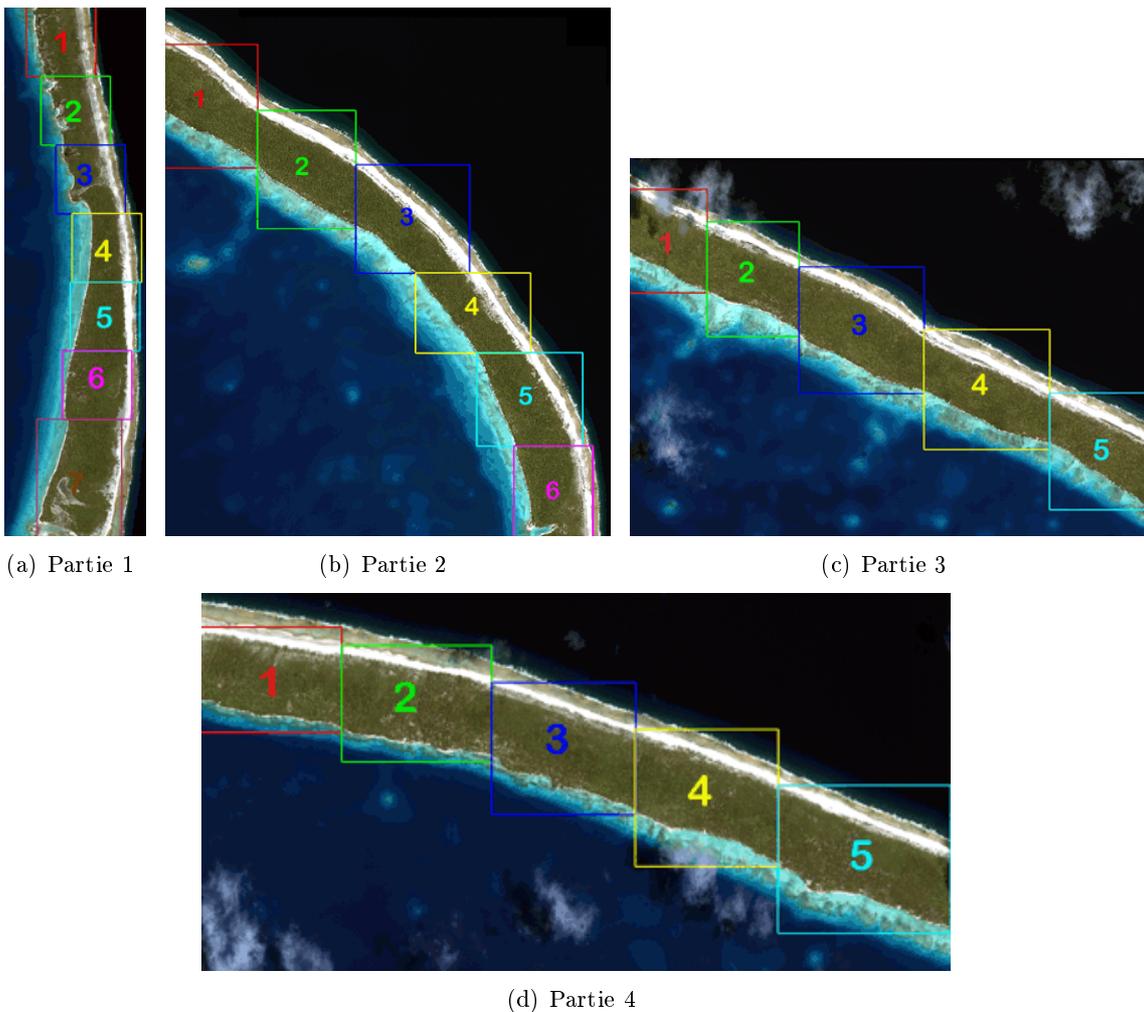


FIG. G.1 – Différentes subdivisions en imagerie du *motu* 13 pour chacune des 4 parties principales.

Les résultats obtenus pour chacune des quatre parties formant le *motu* 13 sont présentés dans le tableau G.1.

Statistiques pour le Motu 13_1						Statistiques pour le Motu 13_2							
Nb cocotiers	8477					Nb cocotiers	11802						
Statistiques des surfaces						Statistiques des surfaces							
moy.	48,16	std	22,12	var	490,16	moy.	49,91	std	22,00	var	484,44		
Statistiques des couleur						Statistiques des couleur							
Bande R	moy.	65,00	std	8,46	var	72,27	Bande R	moy.	57,10	std	7,92	var	64,11
Bande V	moy.	54,94	std	8,55	var	73,55	Bande V	moy.	55,85	std	7,76	var	62,03
Bande B	moy.	31,87	std	5,89	var	35,22	Bande B	moy.	29,71	std	5,11	var	26,61
Partie 1						Partie 2							
Statistiques pour le Motu 13_3						Statistiques pour le Motu 13_4							
Nb cocotiers	14384					Nb cocotiers	13960						
Statistiques des surfaces						Statistiques des surfaces							
moy.	48,15	std	21,51	var	463,12	moy.	51,97	std	22,93	var	527,18		
Statistiques des couleur						Statistiques des couleur							
Bande R	moy.	57,66	std	7,81	var	63,10	Bande R	moy.	53,07	std	7,13	var	50,82
Bande V	moy.	55,95	std	7,57	var	60,56	Bande V	moy.	47,98	std	6,75	var	45,61
Bande B	moy.	30,52	std	5,17	var	27,44	Bande B	moy.	26,96	std	5,09	var	26,07
Partie 3						Partie 4							

TAB. G.1 – Résultats sur les différentes parties du *motu* 13.

G.2 *motu* 16

Pour les mêmes raisons que le *motu* 13, nous avons divisé en plusieurs images plus petites le *motu* principal.

Les résultats obtenus pour chacune des quatre parties formant le *motu* 16 sont présentés dans le tableau G.2.

Statistiques pour le Motu 16_1						Statistiques pour le Motu 16_2							
Nb cocotiers	3295					Nb cocotiers	3984						
Statistiques des surfaces						Statistiques des surfaces							
moy.	57,33	std	24,49	var	599,99	moy.	55,08	std	25,12	var	631,04		
Statistiques des couleur						Statistiques des couleur							
Bande R	moy.	63,92	std	11,81	var	139,50	Bande R	moy.	53,72	std	10,82	var	117,16
Bande V	moy.	52,32	std	11,93	var	142,31	Bande V	moy.	55,66	std	10,64	var	113,11
Bande B	moy.	43,22	std	10,50	var	110,30	Bande B	moy.	38,00	std	8,88	var	78,83
Partie 1						Partie 2							
Statistiques pour le Motu 16_3						Statistiques pour le Motu 16_4							
Nb cocotiers	2907					Nb cocotiers	3385						
Statistiques des surfaces						Statistiques des surfaces							
moy.	56,18	std	25,10	var	629,78	moy.	53,90	std	23,38	var	546,66		
Statistiques des couleur						Statistiques des couleur							
Bande R	moy.	54,28	std	10,29	var	105,82	Bande R	moy.	49,44	std	9,52	var	90,60
Bande V	moy.	49,39	std	10,46	var	109,33	Bande V	moy.	45,90	std	10,17	var	103,33
Bande B	moy.	42,00	std	8,41	var	70,80	Bande B	moy.	39,85	std	8,45	var	71,43
Partie 3						Partie 4							

TAB. G.2 – Résultats sur les différentes parties du *motu* 16

Extraction des informations de texture

Descripteurs de texture du premier ordre

L'histogramme de premier ordre exprime l'occurrence d'une valeur de niveau de gris (N_g) dans un voisinage considéré (une fenêtre). En prenant uniquement les informations des pixels du voisinage, sans connaître les relations entre ces derniers, les informations de texture locales ci-dessous peuvent être calculées :

$$\text{moyenne} = \sum_{b=0}^{N_g-1} b\mathbb{P}(b) = S_M \quad (\text{H.1})$$

$$\text{entropie} = - \sum_{b=0}^{N_g-1} \mathbb{P}(b) \log_2(\mathbb{P}(b)) \quad (\text{H.2})$$

$$\text{écart type} = \left[\sum_{b=0}^{N_g-1} (b - S_M)^2 \mathbb{P}(b) \right]^{\frac{1}{2}} = S_D \quad (\text{H.3})$$

$$\text{coefficient d'asymétrie (skewness)} = \frac{1}{S_D^3} \left[\sum_{b=0}^{L-1} (b - S_M)^3 \mathbb{P}(b) \right] \quad (\text{H.4})$$

où $\mathbb{P}(b) = \frac{N(b)}{F}$ avec F le nombre de pixels dans la fenêtre, $N(b)$ le nombre de pixels de niveau de gris b dans la fenêtre et N_g le nombre de niveaux de gris.

Matrices de cooccurrence

La matrice de cooccurrence P exprime les occurrences de couple de niveau de gris (x_u, x_v) séparés par une distance d et situés sur une droite faisant un angle θ par rapport à l'axe horizontal (la distance d et l'angle θ sont caractérisés par le vecteur directeur $v = (v_x, v_y)$).

Considérons une image I rectangulaire de dimension (N_x, N_y) ayant N_g niveaux de gris. La matrice de cooccurrence P , pour une direction et une distance donnée par le vecteur $v = (v_x, v_y)$ et pour tout couple d'entier (i, j) appartenant à l'ensemble $[0, N_g[$ est définie par :

$$P_{(u,v)} = \sum_{p=1}^{N_x} \sum_{q=1}^{N_y} \mathbb{1}_{I(p,q)=i \text{ et } I(p+u,p+v)=j} \quad (\text{H.5})$$

$P_{(u,v)}(i, j)$ contient donc le nombre d'occurrence des transitions $i \rightarrow j$ et $j \rightarrow i$ dans l'image I entre deux pixels distants de (v_x, v_y) . La matrice P est symétrique et de dimension $N_g \times N_g$. Ces matrices peuvent être normalisées par le nombre total de paires de voisins intervenant dans le calcul afin d'obtenir des probabilités.

Pour montrer l'intérêt de ces matrices, considérons les quatre directions indiquées dans la figure H.1 et calculons les matrices de cooccurrence respectivement à ces quatre directions. La figure H.2 montre les quatre matrices calculées pour une image fictive.

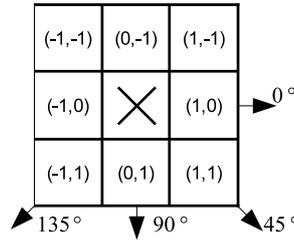


FIG. H.1 – Exemple de distance et d'orientation pour le calcul des matrices de cooccurrence.

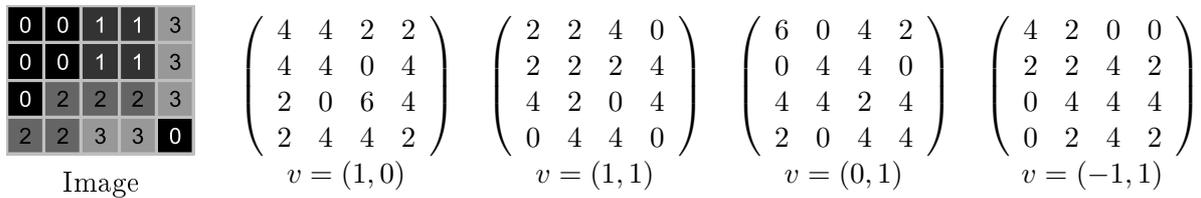


FIG. H.2 – Exemple de calcul de matrices de cooccurrence à partir d'une image de 4×5 avec 4 niveaux de gris.

Prenons le cas de la matrice de cooccurrence calculée pour la direction $v = (1, 0)$. La matrice P indique que le couple $(2, 2)$ est rencontré six fois dans l'image. Le calcul de cette valeur a fait intervenir 32 combinaisons horizontales (16 dans la direction $(1, 0)$ et 16 dans la direction opposée $(-1, 0)$). Au final, le couple $(2, 2)$ apparaît 6 fois parmi 32 combinaisons horizontales possible, ce qui veut dire que l'on a $\frac{6}{32}$ chance de trouver ce couple dans l'image. On divise chacun des termes de la matrice de cooccurrence par le nombre de combinaison horizontales (ici 32), pour obtenir une matrice $P(i, j)$ décrivant la probabilité de rencontrer la transition d'un niveau de gris i vers un niveau de gris j dans l'image.

$$\begin{pmatrix} 4 & 4 & 2 & 2 \\ 4 & 4 & 0 & 4 \\ 2 & 0 & 6 & 4 \\ 2 & 4 & 4 & 2 \end{pmatrix} \text{ en divisant chacun des } \begin{pmatrix} 0,13 & 0,13 & 0,06 & 0,06 \\ 0,13 & 0,13 & 0 & 0,13 \\ 0,06 & 0 & 0,19 & 0,13 \\ 0,06 & 0,13 & 0,13 & 0,06 \end{pmatrix} \text{ termes par 32}$$

On suit la même démarche pour les autres matrices de cooccurrence.

Descripteurs d'Haralick

Haralick définit 14 attributs texturaux calculés à partir de la matrice de cooccurrence P permettant de caractériser les textures. Nous rappelons dans cette annexe ces descripteurs avec les notations suivantes :

- N_g : Nombre de niveau de gris distincts dans l'image,
- \sum_i (resp. \sum_j) : $\sum_{i=1}^{N_g}$ (resp. $\sum_{j=1}^{N_g}$),
- \sum_{ij} : $\sum_i \sum_j$,
- P_{ij} : entrée (i, j) dans la matrice de cooccurrence normalisée,
- $P_x(i) = \sum_j P_{ij}$: entrée i de la matrice de probabilité obtenue en sommant sur les lignes de $P(i, j)$,
- $P_y(j) = \sum_i P_{ij}$,
- $P_{x+y}(k) = \sum_{i+j=k} P_{ij}$, $k = 2, 3, \dots, 2N_g$: Somme sur les diagonales principales,
- $P_{x-y}(k) = \sum_{|i-j|=k} P_{ij}$, $k = 0, 1, \dots, N_g - 1$: somme sur les diagonales secondaires

- Les moyennes : $\mu_i = \sum_{i,j=0}^{N_g-1} iP_{ij}$, $\mu_j = \sum_{i,j=0}^{N_g-1} jP_{ij}$
- Les variances : $\sigma_i^2 = \sum_{i,j=0}^{N_g-1} (i - \mu_i)^2 P_{ij}$, $\sigma_j^2 = \sum_{i,j=0}^{N_g-1} (j - \mu_j)^2 P_{ij}$
- Les écart-type : $\sigma_i = \sqrt{\sigma_i^2}$, $\sigma_j = \sqrt{\sigma_j^2}$

Remarque H.1 *La traduction en français des différents descripteurs peut varier, de ce fait, nous conserverons les noms anglais définis par Haralick.*

1. Angular Second Moment

Ce descripteur mesure l'homogénéité, une forte valeur de f_1 caractérise une texture homogène.

$$f_1 = \sum_{ij} P_{ij}^2 \quad (\text{H.6})$$

2. Contrast

Une forte valeur de f_2 traduit de grandes variations locales.

$$f_2 = \sum_{n=0}^{N_g-1} n^2 P_{x-y}(n) \quad (\text{H.7})$$

3. Correlation

Ceci donne le coefficient de corrélation de la distribution des niveaux de gris. Plus la texture est régulière, plus f_3 est grand.

$$f_3 = \frac{1}{\sigma_x \sigma_y} \left(\sum_{ij} (ij P_{ij}) - \mu_x \mu_y \right) \quad (\text{H.8})$$

où μ_x, μ_y, σ_x et σ_y sont les moyennes et les écart-types de P_x et P_y .

4. Sum of Squares

$$f_4 = \sum_{ij} (i - \mu)^2 p_{ij} \quad (\text{H.9})$$

5. Inverse Difference Moment

$$f_5 = \sum_{ij} \frac{1}{1 + (i - j)^2} P_{ij} \quad (\text{H.10})$$

6. Sum Average

$$f_6 = \sum_{i=2}^{2N_g} iP_{x+y}(i) \quad (\text{H.11})$$

7. Sum Variance

$$f_7 = \sum_{i=2}^{2N_g} (i - f_6)^2 P_{x+y}(i) \quad (\text{H.12})$$

8. Sum Entropy

$$f_8 = - \sum_{i=2}^{2N_g} P_{x+y}(i) \log(P_{x+y}(i)) \quad (\text{H.13})$$

Étant donné que $P_{x+y}(i)$ peut être nulle et que le $\log(0)$ n'est pas défini, il est recommandé d'utiliser le terme $\log(P_{x+y}(i) + \epsilon)$ où ϵ est une constante très petite et positive.

9. Entropy

$$f_9 = - \sum_{ij} \log(P_{ij}) \quad (\text{H.14})$$

10. Difference Variance

$$f_{10} = \sigma^2(P_{x-y}) \quad (\text{H.15})$$

11. Difference Entropy

$$f_{11} = - \sum_i P_{x-y}(i) \log(P_{x-y}(i)) \quad (\text{H.16})$$

12. Information Measures Of Correlation 1

$$f_{12} = \frac{HXY - HXY1}{\max(HX, HY)} \quad (\text{H.17})$$

$$HXY = - \sum_{ij} P_{ij} \log(P_{ij})$$

$$HXY1 = - \sum_{ij} P_{ij} \log(P_x(i)P_y(j))$$

où HX et HY sont les entropies de P_x et P_y .

13. Information Measures Of Correlation 2

$$f_{13} = \sqrt{(1 - \exp[-2(HXY2 - HXY)])} \quad (\text{H.18})$$

$$HXY1 = - \sum_{ij} P_x(i)P_y(j) \log(P_x(i)P_y(j)) \quad (\text{H.19})$$

14. Maximal Correlation Coefficient

$$f_{14} = \sqrt{\text{second largest eigenvalue of } Q} \quad (\text{H.20})$$

où

$$Q_{ij} = \sum_k \frac{P_{ik}P_{jk}}{P_x(i)P_y(k)}$$

Nous avons vu que chacun de ces descripteurs est calculé pour une distance d et une orientation θ donnés. Pour privilégier aucune orientation, on calcule chacun de ces descripteurs pour des matrices de cooccurrence calculées pour des orientations de 0° , 45° , 90° et 135° , puis moyennés.

Annexe I

Estimation des surfaces de cocoteraies sur les atolls

Cette annexe présente les résultats de l'estimation des surfaces de cocoteraies grâce aux SVM et dont les chiffres se trouvent en conclusion du chapitre 9.

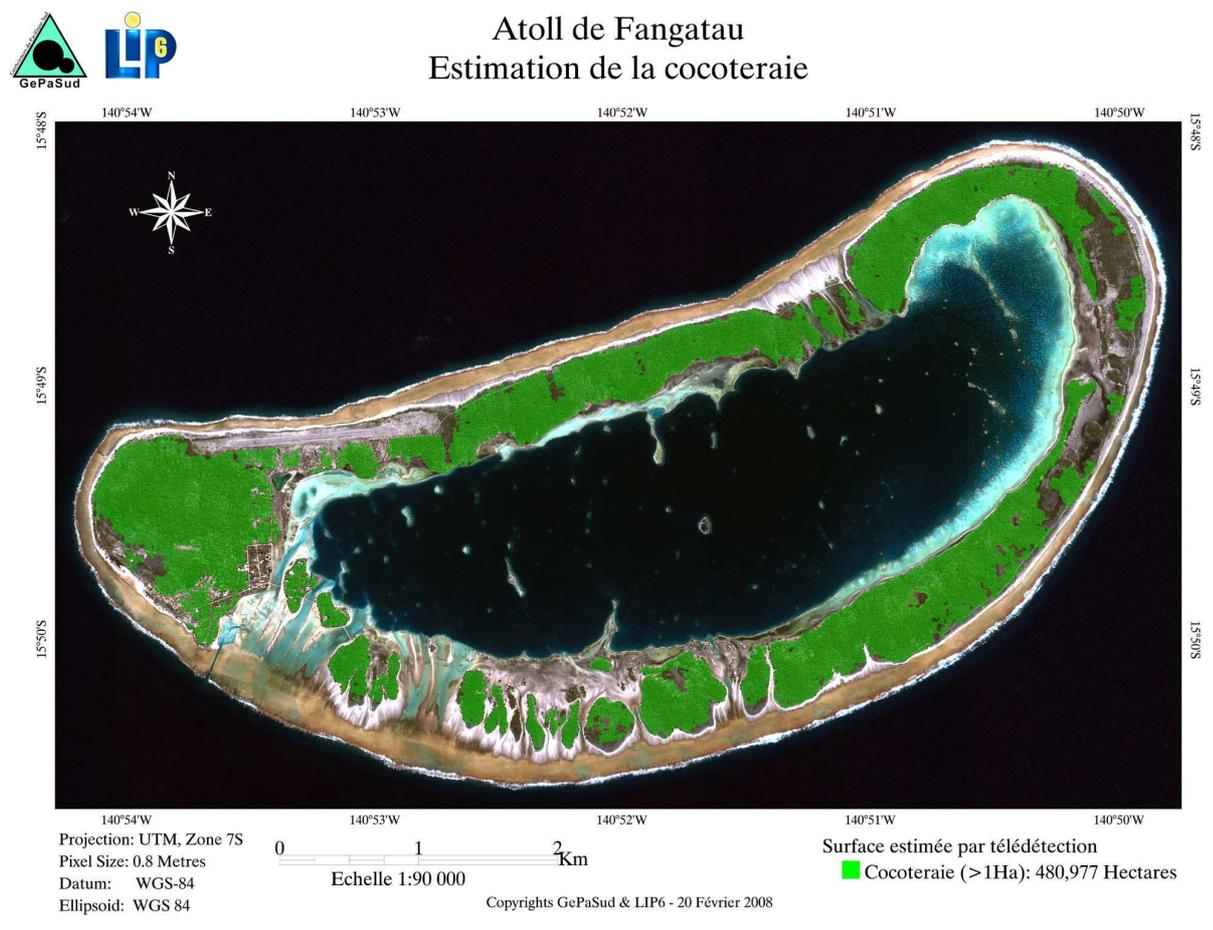


FIG. I.1 – Estimation de la surface de la cocoteraie de l’atoll de Fangatau.



Atoll de Vairaatea
 Estimation de la surface de la cocoteraie



FIG. I.4 – Estimation de la surface de la cocoteraie de l’atoll de Vairaatea.

Validation et extrapolation du nombres d'arbres sur Tikehau

J.1 Protocole des relevés

La figure J.1 présente la fiche terrain utilisée lors des missions. Elle se compose de trois parties :

1. la partie tout en haut contient les informations concernant la placette en cours en mentionnant la date, les coordonnées GPS, le motu sur lequel se trouve la placette, le numéro de placette.
2. La partie centrale
3. La partie tout en bas concerne les cocotiers à l'intérieur de la placette. Pour chaque cocotier, il est décrit :
 - Nb : le numéro d'identification du cocotier,
 - Dhp : diamètre à hauteur de poitrine du cocotier,
 - Ho : hauteur du cocotier,
 - État : état sanitaire (Sain et vigoureux, peu vigoureux, dépérissant et mort),
 - Rectitude : indique la forme du tronc de l'arbre (droit, anguleux, incliné et courbé),
 - Bulbe : forme du bulbe (pas de bulbe, faible et fort),
 - HBulbe : hauteur du bulbe,
 - Nb cica : nombre de cicatrices foliaires pouvant indiquer l'âge du cocotier,
 - Nb reg : le nombre de régimes sur l'arbre,
 - Nb noix : le nombre de noix de coco présentes sur l'arbre,
 - L palme : longueur de la palme,
 - ravageur : présence de parasites.

Pour notre étude, nous nous intéressons principalement aux données concernant le nombre d'arbres par placette avec des contraintes sur la hauteur de chacun d'eux.

J.2 Interpolation des arbres

Dans cette annexe, nous présentons les résultats détaillés de l'extrapolation sur chacun des *motu* constituant l'atoll de Tikehau.

Classes		Vg B	Vg H	Vg E	AF	MF	UMF	SW	N-Veg	Total
Taux moyen		92,70%	64,44%	0,00%	66,76%	72,75%	61,11%	57,04%	0,00%	
Motu 01	Déectés	768	2294	0	0	0	536	0	6	3604
	Extrapolés	828	3560	0	0	0	877	0	6	5271
Motu 02	Déectés	3	503	0	0	0	0	0	0	506
	Extrapolés	3	781	0	0	0	0	0	0	784
Motu 03	Déectés	236	890	0	0	1146	0	0	1	2273
	Extrapolés	255	1381	0	0	1575	0	0	1	3212
Motu 04	Déectés	43	1084	0	0	0	0	0	0	1127
	Extrapolés	46	1682	0	0	0	0	0	0	1729
Motu 05	Déectés	412	2012	0	0	0	0	0	3	2427
	Extrapolés	444	3122	0	0	0	0	0	3	3570
Motu 06	Déectés	808	351	0	424	0	0	0	3	1586
	Extrapolés	872	545	0	635	0	0	0	3	2054
Motu 10	Déectés	82	260	0	585	0	0	0	1	928
	Extrapolés	88	403	0	876	0	0	0	1	1369
Motu 11	Déectés	279	0	6	3689	0	0	0	18	3992
	Extrapolés	301	0	6	5526	0	0	0	18	5850
Motu 12	Déectés	1080	1946	65	2032	0	0	0	19	5142
	Extrapolés	1165	3020	65	3044	0	0	0	19	7312
Motu 13	Déectés	7050	630	41	81094	467	0	188	307	89777
	Extrapolés	7605	978	41	121465	642	0	330	307	131367
Motu 14	Déectés	185	2686	0	0	1038	0	137	10	4056
	Extrapolés	200	4168	0	0	1427	0	240	10	6045
Motu 15	Déectés	0	1342	0	0	0	0	0	80	1422
	Extrapolés	0	2083	0	0	0	0	0	80	2163
Motu 16	Déectés	1390	4451	501	0	20579	798	648	0	28367
	Extrapolés	1499	6907	501	0	28289	1306	1136	0	39638
Motu 17	Déectés	111	236	0	0	0	0	0	55	402
	Extrapolés	120	366	0	0	0	0	0	55	541
Motu 18	Déectés	0	160	0	0	0	0	0	47	207
	Extrapolés	0	248	0	0	0	0	0	47	295
Motu 19	Déectés	207	441	0	0	0	0	0	79	727
	Extrapolés	223	684	0	0	0	0	0	79	987
Motu 20	Déectés	18	655	0	0	0	0	0	89	762
	Extrapolés	19	1016	0	0	0	0	0	89	1125
Motu 22	Déectés	749	889	25	0	0	0	0	126	1789
	Extrapolés	808	1380	25	0	0	0	0	126	2338
Motu 24	Déectés	350	905	12	0	209	0	0	60	1536
	Extrapolés	378	1404	12	0	287	0	0	60	2141

TAB. J.1 – Extrapolation pour chacun des *motu* de l'atoll de Tikehau en fonction des classes.

Bibliographie

- [Anys et al., 1994] Anys, H., Bannari, H., He, D., and Morin, D. (1994). Texture analysis for the mapping of urban areas using airborne multispectral images. (3) :231–245.
- [Asrar et al., 1984] Asrar, G., Fuchs, M., Kanemasu, E., and Hatfield, J. (1984). Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance of wheat. *Agronomy Journal*, 76 :300–306.
- [Baddeley et al., 2005] Baddeley, A., Turner, R., Møller, J., and Hazelton, M. (2005). Residual analysis for spatial point processes (with discussion). *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 67(5) :617–666.
- [Baddeley, 1998] Baddeley, A. J. (1998). Spatial sampling and censoring. In *Stochastic Geometry : Likelihood and Computation, chapter 2*, pages 37–78. Chapman and Hall.
- [Baddeley and van Lieshout, 1993] Baddeley, A. J. and van Lieshout, M. N. M. (1993). Stochastic geometry models in high-level vision. *J. Appl. Stat.*, (20) :233–258.
- [Ball and Hall, 1965] Ball, J. and Hall, D. (1965). ISODATA : A novel method of data analysis and pattern classification. Technical report, Stanford Research Institute, Menlo Park, CA - USA.
- [Baret et al.,] Baret, F., Guyot, G., and Major, D. TSAVI : a vegetation index which minimizes soil brightness effects on lai and apar estimation. In *Proceeding of the 12th Canadian Symposium on Remote Sensing*, volume 3, pages 1355–1358, Vancouver, BC, Canada. IGARSS'90.
- [Bariou et al., 1985] Bariou, R., Lecamus, D., and Le Henaff, F. (1985). *Dossier de Télédétection n°1 : Réponse spectrale des végétaux*. COSTEL, Presses universitaires de Rennes.
- [Bejanin et al., 1994] Bejanin, M., Huertas, A., Medioni, G., and Nevatia, R. (1994). Model validation for change detection. pages 160–167, Los Alamitos, USA. 2nd International IEEE Workshop on Applications of Computer Vision.
- [Béréziat, 1999] Béréziat, D. (1999). *Détection et suivi de structures déformables en mouvement : application à la météorologie*. PhD thesis, Université Paris Sud.
- [Besag, 1977] Besag, J. (1977). Contribution to the discussion of dr ripley's paper. *Journal of the Royal Statistical Society, Series B*, 39 :193–195.
- [Beucher and Lantuéjoul, 1979] Beucher, S. and Lantuéjoul, C. (1979). Use of watersheds in contour detection. In *Proc. International workshop on Image Processing, Real-Time Edge and Motion Detection/Estimation*, Rennes.
- [Burges, 1998] Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2) :121–167.
- [Buschman and Nagel, 1993] Buschman, C. and Nagel, E. (1993). In vivo spectroscopy and internal optics of leaves as basis of remote-sensing of vegetation. *International Journal of Remote Sensing*, pages 711–722.
- [Chang and Lin, 2001] Chang, C.-C. and Lin, C.-J. (2001). *LIBSVM : a library for support vector machines*. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

- [Chantry, 1983] Chantry, G. and Cabannes, Y. W. V. (1983). *Le cocotier : production et mise en oeuvre dans l'habitat*. Ministère de la Coopération, Paris (France). ISBN : 2-11-084847-2.
- [Clark and Evans, 1954] Clark, P. J. and Evans, F. C. (1954). Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, 35.
- [Congalton and Green, 1999] Congalton, R. and Green, K. (1999). *Assessing the Accuracy of Remote Sensing Data : Principles and Practices*. Roca Baton, Lewis.
- [Cook et al., 2001] Cook, M., Peterson, B., Dial, G., Gerlach, F., Hutchins, K., Kudola, R., and Bowen, H. (2001). Ikonos technical performance assesment. Technical report, Space Imaging.
- [Cortes and Vapnik, 1995] Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3) :273–297.
- [Courant and Hilbert, 1953] Courant, R. and Hilbert, D. (1953). *Methods of Mathematical Physics*. Interscience, New York.
- [Cressie, 1993] Cressie, N. (1993). *Statistics for Spatial Data*. Wiley, New York, NY.
- [Culvenor, 2002] Culvenor, D. (2002). Tida : an algorithm for the delineation of tree crowns in high spatial resolution remotely sensed imagery. *Comput. Geosci.*, 28(1) :33–44.
- [Deering et al., 1975] Deering, D., Rouse, J., Hass, R., and Schell, J. (1975). Measuring forage production of grazing units from landsat mss data. In *Proceedings of the 10th International Symposium Remote Sensing Environment*, volume 2, pages 1169–1178.
- [Dial and Grodecki, 2003] Dial, G. and Grodecki, J. (2003). Applications of Ikonos Imagery. Anchorage, Alaska. Space Imaging, ASPRS 2003 Annual Conference Proceeings.
- [Diggle, 1985] Diggle, P. (1985). A kernel method for smoothing point process data. *Applied Statistics*, 34(2) :138–147.
- [Diggle, 1983] Diggle, P. J. (1983). *Statistical Analysis of Spatial Point Patterns*. Academic Press.
- [Field et al., 2002] Field, M., Cochran, S. A., and Evans, K. R. (2002). U.S coral reefs - imperiled national treasure. <http://pubs.usgs.gov/fs/2002/fs025-02>.
- [Fletcher, 1987] Fletcher, R. (1987). *Practical Methods of Optimization*. John Wiley & Sons, New York, 2nd edition.
- [Forgy, 1965] Forgy, E. (1965). Cluster analysis of multivariate date : Efficiency vs. interpretability of classification. *Biometrics*, 21(3) :768.
- [Frost et al., 1982] Frost, V., Stiles, J., Shanmugan, K., and Holtzman, J. (1982). A model for radar images and its application to adaptive digital filtering for multiplicative noise. 4(2) :157–166.
- [GeoEye, 2006] GeoEye (2006). *Ikonos Imagery Products Guide*. GeoEye. Ver 1.5.
- [Gevers and Smeulders, 1999] Gevers, T. and Smeulders, W. M. (1999). Color-based object recognition. *Pattern Recognition*, (32) :453–464.
- [Girard and Girard, 1989] Girard, M. C. and Girard, C. M. (1989). *Téledétection Appliquée : Zones tempérées et intertropicales*. Masson.
- [Gougeon, 1995] Gougeon, F. (1995). A crown-following approach to the automatic delineation of individual tree crowns in high spatial resolution aerial images. *Canadian Journal of Remote Sensing*, 21(3) :274–284.
- [Gougeon et al., 2003] Gougeon, F., Cormier, R., Labrecque, P., Cole, B., Pitt, D., and Leckie, D. (2003). Individual tree crown (itc) delineation on ikonos and quickbird imagery : The cockburn island study. In *Proceeding of 25th Canadian Symposium / 11^{ème} Congrès de l'Association Québécoise de Téledétection*, Montréal, Québec, Canada.
- [Guyot, 1995] Guyot, G. (1995). La réflectance des couverts végétaux. *Photo interprétation*, 33(3) :157–180.
- [Haralick, 1979] Haralick, R. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67 :786–804.

- [Haralick et al., 1973] Haralick, R. M., Dinstein, and Shanmugam, K. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3 :610–621.
- [Holben, 1986] Holben, B. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, 7 :1417–1434.
- [Huang et al., 2004] Huang, J., Xie, W., and Tang, L. (2004). Detection of and compensation for shadows in colored urban aerial images. In *Congress on Intelligent Control and Automation*.
- [Hudson and Ramm, 1987] Hudson, W. and Ramm, C. (1987). Correct formulation of the kappa coefficient of agreement. *Photogrammetric Engineering and Remote Sensing*, 53 :421–422.
- [Huete, 1987] Huete, A. (1987). Soil influence in remote sensed vegetation-canopy spectra. In Elachi, C., editor, *Introduction to the physics and techniques of remote sensing*, pages 107–141. New York, wiley-interscience edition.
- [Huete, 1988] Huete, A. R. (1988). A soil-adjusted vegetation index (savi). *Remote Sensing of Environment*, 25 :295–309.
- [Imaging, 2005] Imaging, S. (2005). IKONOS Planetary Reflectance and Mean Solar Exoatmospheric Irradiance. Rev. 1.
- [Jacq, 2006] Jacq, F. (2006). Evaluation quantitative et qualitative des peuplements de cocotiers sur tikehau. Technical report, Service du Développement Rural - Département FOGER.
- [Jayaraman, 2000] Jayaraman, K. (2000). *A statistical manual for forestry research*. Forestry Statistics and Data Collection - DCA/MISC/01, Bangkok, Thailand.
- [Jensen, 1986] Jensen, J. R. (1986). *Introductory Digital Image Processing : A Remote Sensing Perspective*. Prentice-Hall, New Jersey.
- [Johnson and Zimmer, 1985] Johnson, R. B. and Zimmer, W. J. (1985). A more powerful test for dispersion using distance measurements. *Ecology*, 66(5) :1669–1675.
- [Kailath, 1967] Kailath, T. (1967). The divergence and bhattacharyya distance measures in signal selection. *Communications, IEEE Trans. Communication Theory*, 15(1) :52–60.
- [Karantzalos and Argialas, 2004] Karantzalos, K. and Argialas, D. (2004). Towards automatic olive tree extraction from satellite imagery. *International Society for Photogrammetry and Remote Sensing*.
- [Kaufman and Tanre, 1992] Kaufman, Y. J. and Tanre, D. (1992). Atmospherically resistant vegetation index (arvi) for eos-modis. *IEEE Transactions on Geoscience and Remote Sensing*, (30) :261–270.
- [Kauth and Thomas, 1976] Kauth, R. J. and Thomas, G. S. (1976). The tasseled cap - a graphic description of the spectral-temporal development of agricultural crops as seen by landsat. In *Purdue University*, pages 4–41. West.
- [Knipling, 1970] Knipling, E. B. (1970). Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote Sensing of Environment*, (1) :155–159.
- [Koller et al., 1993] Koller, D., Danilidis, K., and Nagel, H.-H. (1993). Model-based object tracking in monocular image sequences of road traffic scenes. *International Journal of Computer Vision*, 10(3) :257–281.
- [Komura et al., 2002] Komura, R., Kubo, M., and Muramoto, K. (2002). Representation of tree crown by multiple circles in aerial images. In *Proceedings of Geoscience and Remote Sensing Symposium, IGARSS'02*, volume 6, pages 3411–3413. IEEE.
- [Komura et al., 2003] Komura, R., Kubo, M., and Muramoto, K. (2003). Analysis of size of tree crown on aerial image using circle unification. In *SICE Annual Conference*, volume 1, pages 447–451, Fukui, Japan. SICE.
- [Komura et al., 2004] Komura, R., Kubo, M., and Muramoto, K. (2004). Delineation of tree crown in high resolution satellite image using circle expression and watershed algorithm. In *Proceedings of Geoscience and Remote Sensing Symposium, IGARSS'04*, volume 3, pages 1577–1580. IEEE.

- [Kuan et al., 1985] Kuan, D., Sawchuk, A., Strand, T., and Chavel, P. (1985). Adaptive noise smoothing filter for images with signal-dependent noise. *7(2)* :165–177.
- [Larsen, 1998] Larsen, M. (1998). Finding an optimal match window for spruce top detection based on an optical tree model. In *International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry*, Victoria, BC, Canada.
- [Lee, 1986] Lee, J. S. (1986). Speckle suppression and analysis for synthetic aperture radar images. *Optical Engineering*, (25) :636–643.
- [Lillesaeter, 1982] Lillesaeter, O. (1982). Spectral reflectance of partly transmitting leaves : laboratory measurements and mathematical modeling. *Remote Sensing of Environment*, 12 :247–254.
- [Lindeberg, 1994] Lindeberg, T. (1994). Scale-space theory in computer vision, kluwer, dordrecht. *Monograph 1994*.
- [Lopes et al., 1990] Lopes, A., Nezry, E., and Touzi (1990). Adaptive speckle filters and scene heterogeneity. *IEEE Transaction on Geoscience and Remote Sensing*, 28(6) :992–1000.
- [Ludwig and Reynolds, 1988] Ludwig, J. A. and Reynolds, J. F. (1988). *Statistical ecology : a primer on methods & computing*. John Wiley & Sons, Inc., New York, NY, USA.
- [Major et al., 1990] Major, D., Baret, F., and Guyot, G. (1990). A ratio vegetation index adjusted for soil brightness. *Remote Sensing of Environment*, 12.
- [McQueen, 1967] McQueen, J. (1967). Some methods for classification and analysis of the number of clusters. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 281–297.
- [Olofsson, 2002] Olofsson, K. (2002). Detection of single trees in aerial images using template matching. In *ForestSat 2002, Operational tools in forestry using remote sensing techniques*, Heriot Watt University, Edinburgh, Scotland. Monitoring Forest Establishment and Development, Forest Research, Forest Commission, Proceedings CD-ROM, talk FI6.3, session Forestry Inventory 6.
- [Olofsson et al., 2006] Olofsson, K., Wallerman, J., Holmgren, J., and Olsson, H. (2006). Trees species discrimination using z/i dmc imagery and template matching of single trees. *Scandinavian Journal of Forest Research*, 21 :106–110.
- [Otsu, 1979] Otsu, N. (1979). A threshold selection method from grey-level histograms. *SMC*, 9(1) :62–66.
- [Pearson, 1900] Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can reasonable be supposed to have arisen from random sampling. *Philosophical Magazine*, 50 :157–175.
- [Perona and Malik, 1990] Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(7) :629–639.
- [Perrin et al., 2006] Perrin, G., Descombes, X., and Zerubia, J. (2006). Evaluation des ressources forestières à l'aide de processus ponctuels marqués. In *Proc. Reconnaissance des Formes et Intelligence Artificielle (RFIA)*, Tours, France.
- [Pielou, 1977] Pielou, E. (1977). *Mathematical Ecology*. John Wiley & Sons, New York.
- [Polidorio et al., 2003] Polidorio, A. M., Flores, F. C., Imai, N. N., Tommaselli, A. M. G., and Franco, C. (2003). Automatic shadow segmentation in aerial color images. *SIBGRAPI*, 00 :270.
- [Pollock, 1996] Pollock, R. (1996). *The automatic Recognition of Individual trees in Aerial Images of Forests Based on a Synthetic Tree Crown Image Model*. PhD thesis, University of British Columbia, Vancouver, Canada.
- [Puissant et al., 2005] Puissant, A., Hirsch, J., and Weber, C. (2005). The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *International Journal of Remote Sensing*, 26(4) :733–745.
- [Qi et al., 1994] Qi, J., Chehbouni, A., A., H., Kerr, Y., and Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, 48 :119–126.

- [Richards and Jia, 1999] Richards, J. A. and Jia, X. (1999). *Remote Sensing Digital Image Analysis : An Introduction*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- [Richardson and Wiegand, 1977] Richardson, A. J. and Wiegand, C. L. (1977). Distinguishing vegetation from soil background information. (43) :1541–1552.
- [Ricotta and Avena, 1999] Ricotta, C. and Avena, C. (1999). The influence of principal component analysis on the spatial structure of a multispectral dataset. *International Journal of Remote Sensing*, 20(17) :3367–3376.
- [Ripley, 1977] Ripley, B. D. (1977). Modelling spatial patterns. *Journal of the Royal Statistical Society, Series B*, 39 :172–212.
- [Ripley, 1981] Ripley, B. D. (1981). *Spatial Statistics*. Wiley, New York, NY.
- [Rouse et al., 1974] Rouse, J., Haas, R., J.A., S., Deering, D., and J.C., H. (1974). Monitoring the vernal advancement of retrogradation of natural vegetation. NASA/GSFC, Type III, Final Report, Greenbelt. *MD,NASA*, 1 :137.
- [Salvador et al., 2004] Salvador, E., Cavallaro, A., and Ebrahimi, T. (2004). Cast shadow segmentation using invariant color features. *Computer Vision and Image Understanding*, 95(2) :238–259.
- [Sellers, 1985] Sellers, P. (1985). Canopy reflectance, photosynthesis and transpiration. *International Journal of Remote Sensing*, 6 :1335–1372.
- [Shettigara and Sumerling, 1998] Shettigara, V. and Sumerling, G. (1998). Height determination of extended objects using shadows in spot images. *Photogrammetric Engineering and Remote Sensing*, 64(1) :35–44.
- [Stoyan and Stoyan, 1994] Stoyan, D. and Stoyan, H. (1994). Random shapes and point fields, methods of geometrical statistics.
- [Stoyan and Stoyan, 1996] Stoyan, D. and Stoyan, H. (1996). Estimating pair correlation functions of planar cluster processes. *Biometrical Journal*, 38(3) :259–271.
- [Swain, 1978] Swain, P. (1978). *Remote Sensing, the Quantitative Approach*. Mc Graw-Hill.
- [Thirion, 1992] Thirion, J.-P. (1992). Realistic 3d simulation of shapes and shadows for image processing. *CVGIP : Graphical Models and Image Processing*, 54(1) :82–90.
- [Tsai, 2003] Tsai, V. (2003). Automatic shadow detection and radiometric restoration on digital aerial images. *IGARSS*, 2 :732–733.
- [Tsai, 2004] Tsai, V. (2004). A comparative study on shadow compensation of color aerial images in invariant color models. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6) :1661–1671.
- [Tucker and Sellers, 1986] Tucker, C. and Sellers, P. (1986). Satellite remote sensing of primary production. *International Journal of Remote Sensing*, 7(11) :1395–1416.
- [Tucker, 1979] Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, (8) :127–150.
- [van Lieshout, 2000] van Lieshout, M. (2000). *Markov Point Processes and their Applications*. Imperial College Press, London.
- [Vincent and Soille, 1991] Vincent, L. and Soille, P. (1991). Watersheds in digital spaces : An efficient algorithm based on immersion simulations. *IEEE PAMI*, 1991, 13(6) :583–598.
- [Vogelmann, 1990] Vogelmann, J. (1990). Comparison between two vegetation indices for measuring different types of forest damage in north-eastern united states. *International Journal of Remote Sensing*, 11 :2281–2297.
- [Wacker, 1971] Wacker, A. (1971). *The Minimum Distance Approach to Classification*. PhD thesis, Purdue University, West Lafayette.
- [Wang et al., 2004] Wang, L., Gong, P., and Biging, G. (2004). Individual tree-crown delineation and treetop detection in high-spatial-resolution aerial imagery. *Photogrammetric Engineering and Remote Sensing*, 70(3) :351–357.

- [Zarco-Tejada et al., 2000] Zarco-Tejada, P., Miller, J., and Mohammed, G.H. and Noland, T. (December 2000). Chlorophyll fluorescence effects on vegetation apparent reflectance - i. leaf-level measurements and model simulation. *Remote Sensing of Environment*, 74 :582–595(14).
- [Zarco-Tejada et al., 2004] Zarco-Tejada, P. J., Miller, J. R., Morales, A., Berjón, A., and Agüera, J. (2004). Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops. *Remote Sensing of Environment*, 90(4) :463–476.

Table des figures

1.1	Carte de la Polynésie Française.	5
2.1	Séchage préliminaire des noix de coco ouvertes avant détrocage de l'albumen.	11
2.2	Séchage de la chair de coco au soleil.	12
2.3	Évolution de la production d'huile de Coprah	13
2.4	Évolution de l'exportation de Coprah entre 1993 et 2007	13
2.5	Carte de la Polynésie Française.	14
3.1	Schéma de la formation d'un atoll.[Field et al., 2002]	16
3.2	Vue globale de l'atoll de Tikehau.	17
3.3	Fractionnement de l'atoll de Tikehau	17
3.4	Motu principal : localisation des 6 zones de test du motu #16.	18
3.5	Panorama des autres données à notre disposition.	19
3.6	Présentation de l'atoll de Vahitahi.	19
3.7	L'histogramme de l'image 3.6.	20
3.8	Visualisation d'une zone pour les illustrations futures.	20
3.9	Visualisation en détail de la zone du village.	21
3.10	L'atoll de Vahitahi et un zoom de la zone d'étude au Nord-Est.	21
4.1	Exemple de diffusion anisotropique	25
4.2	Apparence des points elliptique en fonction du signe du laplacien [Karantzas and Argialas, 2004].	26
4.3	Application à l'extraction des oliviers d'après [Karantzas and Argialas, 2004].	26
4.4	Modèle synthétique d'arbres [Pollock, 1996].	27
4.5	Voisinage utilisé lors des passages dans les 4 directions	27
4.6	Application des règles de suivi de contour sur des images synthétiques.	28
4.7	Application de la méthode sur une image réelle.	28
4.8	Les 4 directions définissant le domaine de recherche d'un maximum local.	29
4.9	TIDA : Identification des maxima locaux	29
4.10	TIDA : reclassification des pixels	29
4.11	Avant et après le processus d'ajustement [Culvenor, 2002].	29
4.12	Effet de la constante min_k sur le résultat de la croissance de région [Culvenor, 2002].	30
4.13	Position et marques d'une ellipse [Perrin et al., 2006].	30
4.14	Modélisation d'un arbre avec son ombre [Perrin et al., 2006].	31
4.15	Exemples de mouvements simples [Perrin et al., 2006].	31
4.16	Exemple de fusion et de division [Perrin et al., 2006].	31
4.17	Exemple de résultats obtenus sur des données satellites.	32
4.18	Modèle pour un arbre isolé.	32
4.19	Résultat de la détection d'arbres isolés.	32
5.1	Exemple de filtrage Gamma	34
5.2	Exemple de segmentation de texture	34

5.3	Terminologie utilisée pour la LPE.	35
5.4	Processus d'immersion de la LPE.	35
5.5	Processus d'immersion de la LPE avec des marqueurs.	35
5.6	Diagramme de la méthode.	36
5.7	Exemples des différentes natures du sol.	37
5.8	Masquage des zones d'intérêt.	38
5.9	Lissage de l'image.	38
5.10	Marqueurs pour contraindre la LPE	39
5.11	Canopée d'un cocotier inscrit dans un cercle de rayon R	39
5.12	Post-segmentation : élimination des régions aberrantes.	40
5.13	Exemples de données extraites	41
5.14	Exemple de carte de densité locale	41
6.1	Dispositif des placettes en quinconce	45
6.2	Difficulté d'accès dans les zones de corail soulevé.	45
6.3	Schéma des différentes strates	46
6.4	Exemples de différentes strates observées sur le terrain.	46
6.5	Le Motu 25.	48
6.6	Photos du Motu 25.	49
6.7	Résultats en image du <i>motu</i> 25.	49
6.8	Données extraites sur le <i>motu</i> 25	49
6.9	Présentation du grand <i>motu</i> au Nord-Est de l'atoll de Tikehau.	50
6.10	Subdivision des images.	51
6.11	Présentation du <i>motu</i> principal de Tikehau avec les quatre parties traitées.	52
6.12	Exemple de couverture nuageuse	53
6.13	Exemple de traitement séparé pour les zones d'ombres.	54
6.14	Exemple d'arbres isolés exclus du masque final de la végétation haute.	54
6.15	Exemples d'erreurs de segmentation et type d'agglomération.	55
7.1	Transformation de l'image RGB en HSI.	58
7.2	Résultat de la segmentation des ombres (Polidorio)	58
7.3	Image de test pour la méthode.	59
7.4	Les différentes étapes de la méthode de Huang.	60
7.5	Résultat de la segmentation des ombres (Huang)	60
7.6	Configuration de la position du soleil, du capteur et de l'arbre.	61
7.7	Carte des ombres.	62
7.8	Choix des deux lignes d'ombres	62
7.9	Régression linéaire des lignes d'ombres	63
8.1	Principe de la classification par parallélépipèdes.	69
8.2	Illustration de la distribution des densité de probabilité.	71
8.3	Position d'un point x par rapport à un plan P	72
8.4	Distance d'un point x par rapport à un plan P	73
8.5	Illustration du plan médian P	73
8.6	Sensibilité aux vecteurs supports.	75
8.7	Exemple de transformation de données non linéaire	76
9.1	Localisation des différentes classes sur l'image originale.	80
9.2	Histogrammes des neuf classes.	81
9.3	Histogrammes des ensembles d'apprentissage des neuf classes.	81
9.4	Résultat de l'ACP sur l'image RGB	82
9.5	Histogramme des trois composantes après l'ACP.	83
9.6	Différentes organisations spatiales des cocoteraies artificielles.	84
9.7	Descripteurs de texture moyens	84

9.8	Classification par SVM de l'occupation des sols de l'atoll de Vahitahi.	87
10.1	Exemple des différents types plantations.	90
10.2	Résultat de la classification bayésienne et SVM	90
10.3	Évolution de la séparabilité des classes	91
10.4	Performance des classifieurs.	92
10.5	Influence des paramètres P et γ	93
10.6	Influence des paramètres P et γ	93
10.7	Résultat de la classification par SVM	94
10.8	L'atoll de Tikehau et le <i>Motu</i> principal.	94
10.9	Vérité terrain et résultats de la classification par SVM	95
10.10	Carte de densité et position des placettes	96
11.1	Simulation de processus ponctuel de Poisson homogène	100
11.2	Simulation de processus ponctuel de Poisson non-homogène	101
11.3	Estimations de la fonction d'intensité $\lambda(u)$	102
11.4	Méthode de calcul de la fonction $K(r)$	103
11.5	Estimation des fonctions K et L	103
11.6	Atoll de Tikehau et zone d'étude	104
11.7	Les régions étudiées ainsi que leur sous-régions.	104
11.8	Test statistique de Pearson pour les deux sous-régions.	105
11.9	Estimation des fonctions d'intensité λ	106
11.10	Graphe de la fonction $r \mapsto L(r) - r$	106
11.11	Fonctions de corrélation de paire de points $g(r)$	107
11.12	Les histogramme des densités estimées	108
11.13	Seuillages des sous-régions #01 et #04.	108
11.14	Diagramme Quantile-Quantile des classes segmentées de la sous -région #04.	109
12.1	Déplacement de l'ombre	113
A.1	Le satellite Ikonos-2.	117
A.2	Orbite du satellite Ikonos et sa trace au sol.	118
A.3	Réponses Spectrales de la bande pan-chromatique et des bandes multi-spectrales.	119
B.1	Structure physiologique d'une feuille [Girard and Girard, 1989].	122
B.2	Réponse spectrale des végétaux et du sol.	122
B.3	Composition colorée du NDVI calculé sur l'ensemble de la Terre.	123
C.1	Génération du masque de la végétation haute.	127
C.2	Génération du masque de la végétation haute.	128
C.3	Génération du masque de la végétation haute.	129
C.4	Génération du masque de la végétation haute.	130
C.5	Génération du masque de la végétation haute.	131
C.6	Génération du masque de la végétation haute.	132
C.7	Génération du masque de la végétation haute.	133
C.8	Génération du masque de la végétation haute.	134
C.9	Génération du masque de la végétation haute.	135
C.10	Génération du masque de la végétation haute.	136
C.11	Génération du masque de la végétation haute.	137
D.1	Cartes de densité des secteurs de test du motu 16.	139
E.1	Interface de GIMP.	141
E.2	Exemple d'un secteur et du calque en jouant avec l'opacité de ce dernier.	141
E.3	Exemple d'arbres visibles sur les images.	142

G.1	Subdivision du <i>motu</i> 13	147
H.1	Exemple de distance et d'orientation pour le calcul des matrices de cooccurrence.	150
H.2	Exemple de calcul de matrice de cooccurrence	150
I.1	Estimation de la surface de la cocoteraie de l'atoll de Fangatau.	153
I.2	Estimation de la surface de la cocoteraie de l'atoll de Nukutavake.	154
I.3	Estimation de la surface de la cocoteraie de l'atoll de Vahitahi.	154
I.4	Estimation de la surface de la cocoteraie de l'atoll de Vairaatea.	155
J.1	Fiche terrain utilisée par Fred JACQ pour les missions.	158

Liste des tableaux

2.1	Production de Coprah par archipel	12
2.2	Quantités de Coprah entrées en silo en 2007	12
3.1	Liste des nouvelles données recueillies.	18
3.2	Statistiques de l'image 3.6.	20
5.1	Comparaison des statistiques de l'image RGB et de l'ACP.	38
5.2	Correspondance de la surface en fonction du rayon	40
6.1	Résultats obtenus sur chacun des secteurs tests.	44
6.2	Résultat de la vérité photo	44
6.3	Nombre de placettes par secteur	45
6.4	Correspondance des données terrain et densités	47
6.5	Extrapolation des résultats	48
6.6	Résultat du comptage exhaustif	49
6.7	Tableau des résultats sur l'ensemble du <i>motu</i> 13.	51
6.8	Tableau des résultats sur l'ensemble du <i>motu</i> 16.	51
6.9	Résultats de la détection des cocotiers sur tous les <i>motu</i>	53
7.1	Quelques valeurs du fichier de méta données.	62
7.2	Résultats de l'estimation des hauteurs.	63
9.1	Acronyme des différentes classes utilisées.	80
9.2	Séparabilité des classes calculée sur les bandes spectrales RGB.	82
9.3	Les statistiques des composantes principales.	83
9.4	Séparabilité des classes calculée sur les textures.	85
9.5	Séparabilité des classes calculée sur les textures et les bandes spectrales.	85
9.6	Récapitulatif des séparabilité des classes.	85
9.7	Comparaison des classifieurs.	86
9.8	Surfaces de cocoteraies estimées.	88
10.1	Séparabilité des classes.	90
10.2	Méthode de calcul de la performance moyenne	91
10.3	Acronyme des classes utilisées.	94
10.4	Séparabilité des classes.	95
10.5	Exemple de données extraites de la base des relevés terrains.	96
10.6	Taux de détection en fonction des classes.	97
10.7	Extrapolation finale du nombre d'arbres sur l'ensemble de l'atoll de Tikehau.	97
11.1	Quelques statistiques calculées sur les régions et sous-régions d'études.	104
11.2	Détermination des indices de dispersion.	105

11.3	Résumé de chacune des zones segmentées pour les sous-régions #01 et #04.	109
A.1	Caractéristiques du radiomètre du satellite Ikonos 2	118
B.1	Indices de végétation les plus courants	124
C.1	Récapitulatif des étapes de génération des masques pour chacun des motus traités. . .	126
F.1	Correspondance des données terrain du secteur 2	144
F.2	Correspondance des données terrain du secteur 3	144
F.3	Correspondance des données terrain du secteur 4	144
F.4	Correspondance des données terrain du secteur 5	145
F.5	Correspondance des données terrain du secteur 6	145
G.1	Résultats sur les différentes parties du <i>motu</i> 13.	148
G.2	Résultats sur les différentes parties du <i>motu</i> 16	148
J.1	Extrapolation pour chacun des <i>motu</i> de l'atoll de Tikehau en fonction des classes. . . .	159

Résumé

Dans ce manuscrit de thèse sont présentés les travaux effectués sur la caractérisation des cocoteraies de l'archipel des Tuamotu en Polynésie Française à partir d'images satellites Ikonos multispectrales à très haute résolution spatiale.

Ce travail rentre dans le cadre d'un programme de régénération de la cocoteraie des atolls mis en place par le Service du Développement Rural (SDR) de la Polynésie Française. Le recensement complet des cocotiers, l'estimation de la surface occupée par les plantations ainsi que des types plantations disponibles sur les atolls permettent d'avoir un état des lieux de la situation servant de base pour un programme de renouvellement des arbres. De part la situation géographique des atolls, tous ne sont pas directement accessibles par avion et nécessitent de gros moyens financiers et du temps pour organiser les missions sur le terrain. Grâce à l'utilisation d'images satellites à très haute résolution spatiale, dont il est possible d'avoir une résolution de 80 cm au sol du pixel avec le satellite Ikonos, il devient beaucoup plus économique d'effectuer ce travail à distance bien que la connaissance du terrain reste partiellement nécessaire pour valider les résultats.

Ce travail s'articule sur deux parties : la segmentation individuelle des cocotiers afin d'obtenir une estimation du nombre total d'arbres visibles sur les images, puis une caractérisation des types de plantation prenant en compte la densité locale, l'organisation spatiale et la régularité.

Concernant la segmentation individuelle des cocotiers, un état de l'art non exhaustif des méthodes utilisées pour la détection des couronnes des arbres est présenté. Puis nous détaillons une méthode de dénombrement des cocotiers à partir d'images Ikonos dépourvues de bande proche infrarouge. L'absence de la bande proche infrarouge est compensée par un procédé judicieux et original de classification bayésienne qui permet d'éliminer les zones non végétales. Un algorithme de segmentation par LPE est initié à partir de marqueurs judicieusement choisis. Enfin une dernière étape élimine les zones segmentées non pertinentes. Cette méthode peut également traiter des zones sous ombres portées des nuages en modifiant de façon adéquate le procédé de masquage des zones de végétation. Enfin, un travail de validation sur le terrain a été réalisé.

Le second volet concerne la caractérisation des plantations. Après avoir présenté un état de l'art des méthodes de classification, les performances de ces dernières sont alors comparées pour un problème d'occupation des sols. Puis, nous mettons en évidence la nécessité d'ajouter des informations de texture pour pouvoir caractériser les types de plantations avec des méthodes de classification. Une étude expérimentale permet de conclure que les classifieurs de type SVM ont les meilleures performances en terme de qualité de classification. Néanmoins, ces approches sont coûteuses en ressources informatiques car elles classifient chaque pixel d'images de très grande taille. Pour résoudre ce problème, nous proposons une approche duale, qui classe les types de cocoteraies à partir de la répartition spatiale des coordonnées de chaque cocotier. Cette méthode repose sur une modélisation de type processus ponctuel de Poisson.

Tous ces algorithmes ont été appliqués avec succès sur de grandes quantités de données et font maintenant l'objet d'une exploitation par le SDR.