

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE  
UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À  
L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

COMME EXIGENCE PARTIELLE  
À L'OBTENTION DE LA  
MAÎTRISE EN GÉNIE DE LA PRODUCTION AUTOMATISÉE

M. Ing.

PAR  
Jonathan BOUCHARD

MÉTHODES DE VISION ET D'INTELLIGENCE ARTIFICIELLES POUR LA  
RECONNAISSANCE DE SPÉCIMENS CORALLIENS

MONTRÉAL, LE 6 AVRIL 2011



, Jonathan Bouchard, 2011



Cette licence [Creative Commons](https://creativecommons.org/licenses/by-nc-nd/4.0/) signifie qu'il est permis de diffuser, d'imprimer ou de sauvegarder sur un autre support une partie ou la totalité de cette œuvre à condition de mentionner l'auteur, que ces utilisations soient faites à des fins non commerciales et que le contenu de l'œuvre n'ait pas été modifié.

**PRÉSENTATION DU JURY**

CE MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE :

M. Jacques-André Landry, directeur de mémoire  
Département de génie de la production automatisée à l'École de technologie supérieure

M. Robert Sabourin, président du jury  
Département de génie de la production automatisée à l'École de technologie supérieure

M. Robert Hausler, membre du jury  
Département de génie de la construction à l'École de technologie supérieure

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 25 MARS 2011

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE



## REMERCIEMENTS

Tout d'abord, j'aimerais remercier mon directeur de recherche Jacques-André Landry de m'avoir proposé ce travail passionnant qui allie deux de mes nombreuses passions, soit l'ingénierie et la nature. Également, pour son humanité, pour m'avoir permis sans hésiter, tout au long du parcours, de profiter au maximum de la vie étudiante en m'impliquant dans plusieurs regroupements qui me tiennent à cœur.

Par ailleurs, je tiens à remercier ces gens merveilleux et plein d'énergie qui m'ont accompagné dans cette implication, Martin et Pascale, des Services aux Étudiants, sans qui mon parcours à l'ÉTS n'aurait pas été aussi agréable. Aussi, mes collègues du LIVIA, qui par leur générosité m'ont fait profiter de conseils judicieux, tant sur le plan académique que personnel.

Enfin, j'aimerais remercier l'École de technologie supérieure, pour son appui dans tous mes projets. D'abord pour ses gens. Traiter avec le personnel de l'ÉTS a été une expérience des plus agréables. Merci donc à ces personnes compétentes qui m'ont aidé dans la démarche, que ce soit pour du support informatique ou à la bibliothèque. De plus, merci pour le financement accordé à ce projet de maîtrise, sans quoi je n'aurais pu m'y consacrer totalement.

Merci également à la communauté du logiciel libre, sans quoi aucune des expérimentations présentées n'aurait été possible. Et enfin, mille mercis à ma famille et à mes amis, pour leur soutien et leurs conseils pleins de bon sens! Et plus particulièrement à Frédéric Madore, cet ami, ce personnage rassembleur et tellement coloré, qui à notre grand regret nous a quittés en 2010. Merci d'avoir été présents.



# MÉTHODES DE VISION ET D'INTELLIGENCE ARTIFICIELLES POUR LA RECONNAISSANCE DE SPÉCIMENS CORALLIENS

Jonathan BOUCHARD

## RÉSUMÉ

Ce mémoire traite de l'automatisation de l'extraction de données nécessaires à la caractérisation de la biodiversité d'espèces benthiques, une problématique environnementale d'actualité. L'impact des activités humaines sur la faune marine est une préoccupation grandissante. Les eaux des Philippines, qualifiées par plusieurs écologistes comme étant le « berceau de la biodiversité marine », sont comme tant d'autres sous la menace constante des changements climatiques, de la pêche invasive et de la pollution côtière. Notre sujet d'intérêt, les récifs coralliens, sont une des principales victimes de ces perturbations humaines. Le phénomène est largement reconnu, mais malheureusement aucune évaluation quantitative des dommages causés à la biodiversité ne peut être obtenue facilement. Les écologistes ont besoin d'outils, permettant de traiter en lot les données recueillies sur le terrain pour conclure sur le véritable impact de l'homme sur cet écosystème fragile.

Dans ce contexte, le but de l'étude est de fournir une implémentation logicielle capable d'automatiser la détection et l'identification de spécimens coralliens dans une banque de photographies sous-marines. Pour ce faire, deux approches distinctes ont été mises à l'essai : l'identification des coraux en fonction de leur forme de croissance (la représentation par *formes benthiques*) et l'identification précise de l'espèce des spécimens (la représentation *taxinomique*). Pour l'identification par formes benthiques, divers outils de vision artificielle, tels les descripteurs issus des *matrices de cooccurrence des tons de gris* (GLCM), des *motifs binaires locaux* (LBP), de l'*histogramme* des tons et de la *transformée de Fourier* ont été mis à l'essai. Ces descripteurs ont été utilisés conjointement à divers algorithmes de reconnaissance de formes tels le classificateur *bayésien*, la méthode des *plus proches voisins*, l'*arbre de classification C4.5*, le *séparateur à vaste marge* (SVM) et le *perceptron multicouches* qui ont été comparés dans ce contexte de classification par formes benthiques.

Dans un deuxième temps, l'expérimentation a été reprise sur le problème de classification taxinomique. Au cours de cette étude, la problématique de segmentation automatisée des spécimens de corail a été abordée. Une présegmentation par *laplacien du gaussien* et partage des eaux, suivie d'une classification des segments avec un SVM optimisé, utilisant les descripteurs développés précédemment, se sont avérées une solution efficace à ce problème. Par la suite, les segments appartenant au corail (et non à l'arrière plan) ont été classifiés à l'aide d'un autre SVM entraîné à cet effet. Les résultats obtenus démontrent qu'il est possible d'implémenter un tel système de reconnaissance automatisée du benthos. Toutefois, plusieurs conditions devront être respectées pour en assurer le succès : (1) pour être robuste, le système de segmentation et de classification de corail doit être entraîné avec un nombre suffisant d'échantillons et (2) le protocole d'acquisition des données et d'extraction de descripteurs doit être bien défini pour maximiser les performances.





# COMPUTER VISION AND DATA MINING METHODS APPLIED TO THE RECOGNITION OF CORAL SPECIMENS

Jonathan BOUCHARD

## ABSTRACT

This thesis deals with the automated data extraction process needed to characterize benthic species biodiversity and its evolution: a modern environmental issue. Human impact on marine life is a growing concern. The Philippines seas, described by many environmentalists as the "cradle of marine biodiversity" are like too many others under constant threats by climate change, invasive fisheries and coastal pollution. Our topic of interest: the coral reefs are one of the main victims of these human perturbations. The phenomenon is widely recognized, but unfortunately no quantitative evaluation of the damage to this ecosystem and its biodiversity can be obtained easily. Environmentalists are in need of automated tools, capable of dealing with the enormous amount data collected undersea, to conclude on the real human impact on the fragile coral ecosystem.

The aim of this study is to provide a software implementation able to automate the detection and identification of coral specimens within a vast underwater photographic database. To accomplish this task, two distinct approaches have been tested: identification of corals based on their growth form (*benthic forms identification*) and accurate identification of species for each specimen (*taxonomic identification*). To complete benthic forms identification, various computer vision tools, such as descriptors extracted from the *gray-level cooccurrence matrix* (GLCM), the *local binary patterns* (LBP), the *color channels histograms* and the *Fourier transform* were tested. These descriptors were used in conjunction with various machine learning algorithms such as the *bayesian classifier*, the method of *nearest neighbors*, the *C4.5 classification tree*, the *support vector machine* (SVM) and the *multilayer perceptron* (MLP) which were compared in this context of classification based on benthic forms.

Following the study on benthic forms identification, a similar experiment was conducted on the taxonomic classification problem. In this part of the study, the problem of automated segmentation of coral specimens was addressed. A combination of the *Laplacian of Gaussian* and *watershed* segmentation techniques was used to perform image tessellation. This tessellation was followed by a classification of resulting segments with an optimized SVM, using the previously developed descriptors. This method proved to be an effective solution to the segmentation problem. Subsequently, coral segments were classified using a different SVM trained for this specific purpose. The results show that it is possible to implement such an automated benthos recognition system. However, several conditions must be met to ensure success: (1) to be robust, the segmentation and classification systems should be driven with a sufficient number of samples and (2) the data acquisition and the feature extraction protocols must be well defined to maximize performance.



## TABLE DES MATIÈRES

	Page
INTRODUCTION .....	1
CHAPITRE 1 ÉTAT DE L'ART .....	9
1.1 Les recherches précédentes sur la reconnaissance automatisée de corail .....	9
1.1.1 Les recherches basées sur la télédétection .....	10
1.1.2 Les recherches basées sur l'imagerie sous-marine .....	11
1.1.3 Discussion sur les recherches précédentes .....	15
1.2 Les travaux antérieurs à ce mémoire .....	16
1.2.1 L'intelligence artificielle pour la classification d'objets biologiques .....	16
1.2.2 L'optimisation de la taille des régions d'intérêt .....	20
1.3 Les sujets abordés en vision artificielle et reconnaissance de formes .....	21
1.3.1 L'acquisition des images et les prétraitements .....	22
1.3.1.1 La balance des couleurs .....	22
1.3.1.2 L'ajustement de la plage dynamique .....	23
1.3.1.3 La spectrométrie .....	24
1.3.1.4 La stéréovision .....	25
1.3.2 L'échantillonnage .....	25
1.3.2.1 La validation simple .....	26
1.3.2.2 La stratification .....	26
1.3.2.3 La validation croisée .....	27
1.3.3 La segmentation .....	27
1.3.3.1 Les fondements de la vision artificielle .....	28
1.3.3.2 Le seuillage .....	30
1.3.3.3 Le laplacien du gaussien .....	32
1.3.3.4 Le partage des eaux .....	33
1.3.3.5 Les méthodes avancées .....	34
1.3.4 L'extraction de caractéristiques .....	41
1.3.4.1 Le traitement de l'histogramme des tons .....	41
1.3.4.2 La matrice de cooccurrence des tons de gris .....	43
1.3.4.3 Les motifs binaires locaux .....	47
1.3.4.4 La transformée de Fourier .....	49
1.3.4.5 La réduction de dimensionnalité .....	51
1.3.5 L'optimisation de paramètres et l'entraînement .....	52
1.3.5.1 Le classificateur naïf de Bayes .....	53
1.3.5.2 L'algorithme des $k$ plus proches voisins .....	53
1.3.5.3 L'arbre de classification C4.5 .....	54
1.3.5.4 Le réseau de neurones à base radiale .....	55
1.3.5.5 Le séparateur à vaste marge .....	56
1.3.5.6 Le perceptron multicouche .....	58

1.3.6	La généralisation.....	60
1.3.6.1	La fusion des décisions.....	60
1.3.6.2	Le mécanisme de rejet.....	61
1.3.6.3	Les matrices de confusion.....	63
1.3.6.4	Les intervalles de confiance.....	65
1.4	Discussion sur la revue de littérature.....	66
CHAPITRE 2 APPROCHE DE DESCRIPTION PAR FORMES BENTHIQUES.....		67
2.1	Introduction aux formes benthiques.....	68
2.2	L'acquisition des images et les prétraitements.....	72
2.3	La base de données.....	73
2.4	La segmentation.....	75
2.5	Les expérimentations sur l'approche de description par formes benthiques.....	76
2.5.1	L'optimisation de la représentation des données.....	76
2.5.1.1	Méthodologie.....	76
2.5.1.2	Résultats expérimentaux.....	79
2.5.1.3	Discussion.....	80
2.5.2	La restructuration de la base de données des Philippines.....	80
2.5.2.1	Méthodologie et résultats expérimentaux.....	81
2.5.2.2	Discussion.....	84
2.5.3	L'implémentation d'un système étagé avec mécanisme de rejet.....	84
2.5.3.1	Méthodologie.....	85
2.5.3.2	Résultats expérimentaux.....	87
2.5.3.3	Discussion.....	87
2.5.4	L'implémentation d'un système d'identification par détecteurs.....	88
2.5.4.1	Méthodologie.....	88
2.5.4.2	Résultats expérimentaux et discussion.....	90
2.6	Discussion générale sur l'approche descriptive par formes benthiques.....	91
CHAPITRE 3 APPROCHE PAR DESCRIPTION TAXINOMIQUE COMPLÈTE.....		93
3.1	Introduction à la taxinomie des coraux.....	94
3.2	L'acquisition des images et la base de données.....	98
3.2.1	Extraction des images de <i>Coral ID</i> .....	99
3.2.2	Sélection des images de <i>Coral ID</i> .....	100
3.3	L'échantillonnage des données et les méthodes de validation.....	102
3.3.1	L'échantillonnage des données.....	102
3.3.1.1	Évaluation de la performance en segmentation.....	104
3.3.1.2	Évaluation de la capacité de classification considérant deux métaclasses.....	104
3.3.1.3	Évaluation de l'effet du nombre de classes sur les performances .....	104
3.3.1.4	Évaluation de l'effet de la quantité de données par classe sur les performances.....	105
3.3.2	La méthode de validation.....	106

3.4	La segmentation des spécimens .....	108
3.4.1	La segmentation manuelle .....	109
3.4.2	La segmentation automatique supervisée .....	118
3.4.2.1	La présegmentation non supervisée .....	119
3.4.2.2	L'extraction des caractéristiques pour la segmentation .....	124
3.4.2.3	L'entraînement d'un algorithme de RDF pour la segmentation .....	125
3.5	L'extraction des caractéristiques pour la classification .....	126
3.5.1	Les descripteurs de morphologie .....	127
3.5.2	Les descripteurs de texture .....	128
3.5.3	Les descripteurs de couleur .....	130
3.5.4	Divergences relatives aux travaux de Levasseur .....	131
3.6	L'entraînement des classificateurs et l'optimisation des métaparamètres .....	131
3.6.1	L'optimisation des métaparamètres .....	133
3.6.2	L'entraînement des algorithmes de RDF .....	135
3.7	La généralisation .....	136
3.8	Les résultats expérimentaux et leur interprétation .....	136
3.8.1	Expérience sur la segmentation automatisée .....	137
3.8.1.1	Analyse de la présegmentation non supervisée .....	137
3.8.1.2	Évaluation de la performance en segmentation .....	139
3.8.2	Expériences sur la classification automatisée .....	144
3.8.2.1	Évaluation de la capacité de classification en considérant deux métaclasse .....	145
3.8.2.2	Évaluation de l'effet du nombre de classes sur les performances .....	146
3.8.2.3	Évaluation de l'effet de la quantité de données par classe sur les performances .....	148
3.9	Discussion générale sur l'approche par description taxinomique complète .....	150
	CONCLUSION .....	153
	RECOMMANDATIONS .....	157
	LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES .....	161



## LISTE DES TABLEAUX

	Page
Tableau 0.1	Historique du projet .....7
Tableau 1.1	Les descripteurs de morphologie .....17
Tableau 1.2	Les descripteurs de texture.....18
Tableau 1.3	Les descripteurs de couleur.....19
Tableau 1.4	Quelques masques de corrélation pour le filtrage.....29
Tableau 1.5	Exemple de matrice de confusion .....63
Tableau 2.1	La codification des formes benthiques .....69
Tableau 2.2	Les codes benthiques supplémentaires .....75
Tableau 2.3	Les échantillonnages testés par l'expérimentation .....77
Tableau 2.4	Les métaparamètres généralement recommandés.....78
Tableau 2.5	Taux de classification moyens reliés aux différents échantillonnages .....79
Tableau 2.6	Répartition des échantillons parmi les classes de la base de données .....82
Tableau 2.7	Statistiques sur la restructuration de la base de données des Philippines .....83
Tableau 2.8	Résultats obtenus sur la base de données des Philippines restructurée .....84
Tableau 2.9	Résultats obtenus sur Philippines avec un système étagé.....87
Tableau 2.10	Métaparamètres pour le système d'identification par détecteurs.....90
Tableau 3.1	Les espèces considérées par l'échantillonnage .....103
Tableau 3.2	Groupes de classes échantillonnés .....105
Tableau 3.3	Taux de classification bruts de segmentation .....139
Tableau 3.4	Matrices de confusion de la segmentation automatisée .....140
Tableau 3.5	Taux de classification bruts de classification (métaclasses).....145
Tableau 3.6	Matrices de confusion de la classification par métaclasses .....146





## LISTE DES FIGURES

	Page
Figure 0.1	Exemple de faible variabilité entre deux espèces. ....3
Figure 0.2	Deux spécimens très différents d' <i>acropora cytherea</i> . ....5
Figure 0.3	Le projet dans son contexte.....6
Figure 1.1	Schéma d'un système de reconnaissance de formes par vision artificielle. ...21
Figure 1.2	Corrélation d'une image de 3 bits par un masque moyennant de taille 3. ....29
Figure 1.3	Exemple de seuillage. ....31
Figure 1.4	L'algorithme d'Otsu.....31
Figure 1.5	Exemple de filtrage LoG.....32
Figure 1.6	Exemple de segmentation par partage des eaux. ....33
Figure 1.7	Exemple de matrice de cooccurrence avec $d = 1$ et $\theta = 0^\circ$ .....44
Figure 1.8	Exemple de filtrage par motifs binaires locaux. ....48
Figure 1.9	Magnitude de la transformée de Fourier.....50
Figure 1.10	Algorithme de la rétropropagation de l'erreur. ....59
Figure 1.11	Algorithme d'ajustement des seuils de Chow.....61
Figure 1.12	Algorithme d'ajustement des seuils de Fumera. ....62
Figure 2.1	Exemples de formes benthiques communes. ....70
Figure 2.2	Divers montages pour la photographie sous-marine.....73
Figure 2.3	Les entités de la base de données des Philippines. ....74
Figure 2.4	Le système étagé avec mécanisme de rejet.....86
Figure 2.5	Le système d'identification par détecteurs. ....89
Figure 3.1	Un spécimen d' <i>acropora hyacinthus</i> . ....96
Figure 3.2	Répartition des espèces en fonction de la quantité d'images de <i>Coral ID</i> . .101
Figure 3.3	Outil de segmentation manuel (état initial).....110

## XVIII

Figure 3.4	Image originale de <i>Coral ID</i> . .....	111
Figure 3.5	Les outils, modes et contrôles de l’outil de segmentation. ....	112
Figure 3.6	Outil de segmentation (configuration de la baguette magique). ....	113
Figure 3.7	Outil de segmentation (après la sélection). ....	114
Figure 3.8	Outil de segmentation (affinement de la sélection). ....	115
Figure 3.9	Outil de segmentation (utilisation du lasso). ....	116
Figure 3.10	Outil de segmentation (résultat final). ....	116
Figure 3.11	Image segmentée par partage des eaux sur le gradient LoG. ....	121
Figure 3.12	Tranche centrale du filtre LoG bidimensionnel (taille 51, $\sigma = 51/8$ ). ....	122
Figure 3.13	Algorithmes heuristiques de fusion des segments. ....	123
Figure 3.14	Image segmentée après fusion des segments. ....	123
Figure 3.15	Topologie de l’espace d’optimisation pour le SVM de segmentation. ....	135
Figure 3.16	Distribution de l’aire des segments de corail issus de la présegmentation. .	138
Figure 3.17	Nombre de segments de corail en fonction de leur classe. ....	138
Figure 3.18	Superficie totale des segments de corail en fonction de leur classe. ....	139
Figure 3.19	Exemple de bonne segmentation ( <i>acropora nobilis 52-06E</i> ). ....	141
Figure 3.20	Exemple de bonne segmentation ( <i>acropora nobilis 52-01M</i> ). ....	142
Figure 3.21	Exemple de bonne segmentation ( <i>acropora glauca 31-01M</i> ). ....	142
Figure 3.22	Exemple de mauvaise segmentation ( <i>cycloseris cyclolites 122-09</i> ). ....	143
Figure 3.23	Exemple de mauvaise segmentation ( <i>stylophora pistillata 345-10M</i> ). ....	143
Figure 3.24	Exemple de mauvaise segmentation ( <i>plerogyra sinuosa 305-07</i> ). ....	144
Figure 3.25	Taux de classification en fonction du nombre de classes. ....	147
Figure 3.26	Taux de classification en fonction de la surface considérée pour le SVM. .	149
Figure 3.27	Taux de classification en fonction de la surface considéré pour le kPPV. .	149

## LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

ACP	Analyse en composantes principales
ADN	Acide désoxyribonucléique
AIMS	<i>Australian Institute of Marine Science</i>
ARFF	<i>Attribute-relation file format</i>
C4.5	Arbre de classification C4.5
CSV	<i>Comma separated values</i> – Valeurs séparées par des virgules
DoG	<i>Difference of Gaussians</i> – différence du gaussien (méthode de vision)
GPS	<i>Global positioning system</i> – Système de positionnement global
HSV	<i>Hue, saturation, value</i> – Teinte, saturation, valeur (espace colorimétrique)
IA	Intelligence artificielle
ISO	<i>International Organisation for Standardisation</i>
JPEG	<i>Joint Photographic Experts Group</i> (format d'image compressée avec perte)
kPPV	k plus proches voisins
Lab	Luminance et chromacité a et b (espace colorimétrique)
LBP	<i>Local binary patterns</i> – motifs binaires locaux
LoG	<i>Laplacian Of Gaussian</i> – laplacien du gaussien (méthode de vision)
Mo	Méga-octets
NB	Algorithme naïf de Bayes
ONU	Organisation des Nations Unies
OoO	OpenOffice.org
PDF	<i>Probability density function</i> – Fonction de densité de probabilité
PM	Perceptron multicouche

XX

PNG	<i>Portable network graphics</i> (format d'image compressée avec ou sans perte)
RBF	<i>Radial basis function</i> – fonction à base radiale
RBR	Réseau de neurone à base radiale
RDF	Reconnaissance de formes
ROC	<i>Receiver operator characteristic</i>
RVB	Rouge, vert, bleu (espace colorimétrique)
SVM	Séparateur à vaste marge, <i>Support vector machine</i>
VA	Vision artificielle
VSA	Véhicule sous-marin autonome (traduit de <i>autonomous underwater vehicle</i> )

## **LISTE DES SYMBOLES ET UNITÉS DE MESURE**

GHz	Gigahertz ( $1 \times 10^9$ Hertz)
nm	Nanomètre



## INTRODUCTION

C'est en décembre 2006, que « profondément préoccupée par l'appauvrissement continu de la diversité biologique », l'Organisation des Nations Unies (ONU) a déclaré l'année 2010 « Année internationale de la biodiversité ». Cette déclaration s'est accompagnée d'une multitude d'objectifs afin d'encourager un mouvement planétaire vers la préservation des écosystèmes pour ainsi limiter les « incidences sociales, économiques, écologiques et culturelles » de l'appauvrissement de la biodiversité (ONU, 19 janvier 2007). Par ailleurs, un des vastes volets de ce mouvement est la protection des milieux aquatiques et des organismes vivants qui y habitent. Ces organismes aquatiques sont la principale énergie motrice du cycle du carbone. À eux seuls, ils sont responsables d'environ 50 % du renouvellement de l'oxygène de la planète, et seraient même à l'origine de la formation de l'atmosphère terrestre (Bourque, 2009; Holland, 2006; Roach, 2004).

En dépit de leur rôle crucial dans la chaîne alimentaire et dans le renouvellement de l'oxygène, les organismes marins sont sous la menace continue des activités humaines. Changements climatiques, pollution, déversements pétroliers, surpêche sont quelques exemples des problèmes engendrés par l'homme qui portent atteinte à l'écosystème marin. Daniel G. Boyce de l'Université Dalhousie d'Halifax déclare que le réchauffement de la planète est la cause du déclin du plancton végétal (ultime base de la chaîne alimentaire) à un rythme de près d'un pour cent par année (Boyce, Lewis et Worm, 2010). À ce sujet, « les données recueillies dans l'hémisphère nord laissent penser que le phytoplancton a diminué d'environ 40 % depuis les années 1950 » (Radio-Canada, 2010). Et pourtant, toujours selon Boyce, la température moyenne à la surface des océans du globe n'a augmenté que de quelques degrés (1 à 2 °C) depuis 1950.

L'exemple précédent démontre qu'il suffit de peu pour perturber cet écosystème. En effet, la survie des organismes aquatiques dépend intimement de complexes relations symbiotiques : interactions qui sont malheureusement menacées lorsque la biodiversité est atteinte par les activités humaines. Par exemple, quiconque se rappelle avoir admiré une photographie d'un

récif, telle la Grande Barrière de corail d'Australie, est capable d'identifier le corail simplement en localisant les zones où les couleurs sont les plus saturées. Les couleurs vives du corail sont en fait dues à leur relation symbiotique avec une algue très colorée, nommée zooxanthelle, qui utilise les structures coralliennes comme abri et le gaz carbonique libéré par les coraux pour survivre. En retour, elles produisent divers nutriments, par photosynthèse, qui sont source d'énergie pour leur hôte et libèrent de l'oxygène. Lorsque cette relation symbiotique est interrompue par un stress externe, le corail expulse l'algue cellulaire et perd ses couleurs vives. Ce phénomène est appelé le blanchiment du corail (*coral bleaching*) et s'accompagne généralement de la mort du spécimen atteint (Veron, 2000).

Sans signifier que leurs actes sont directement liés avec la déclaration de l'ONU, aux Philippines, où les mers environnantes possèdent la plus grande diversité de vie aquatique présente sur terre, une multitude d'organisations militent dans le but de ralentir le rythme de l'appauvrissement de la biodiversité marine (Wilkinson, 2008). Toutefois, les causes de cet appauvrissement ne sont pas toutes évidentes. Pour bien comprendre le phénomène, diriger et justifier les actions à entreprendre, il est impératif de faire l'acquisition et l'analyse de données quantitatives. C'est dans ce domaine que l'ingénierie de la production automatisée peut être utilisée pour accélérer le travail des biologistes et écologistes qui effectuent le suivi de l'évolution de la biodiversité. Ainsi, tout comme les organismes vivants qui doivent être protégés, les biologistes et les ingénieurs travailleront en symbiose pour assurer leur pérennité.

Plus précisément, une équipe de chercheurs de l'Université De La Salle (Manille, Philippines), sous la supervision du Dr Wilfredo Ybañez Licuanan, étudie l'évolution temporelle et spatiale de l'intégrité écologique des récifs coralliens (Cleto L. *et al.*, 2004). Pour arriver à leurs fins, des milliers de photographies sous-marines doivent être acquises, chaque année, et finalement être analysées en vue d'effectuer un suivi de la biodiversité. À cette fin, diverses méthodes pour évaluer la diversité biologique sont utilisées. Toutes ces méthodes possèdent un tronc commun; elles étudient l'hétérogénéité spatiale en mettant en relation deux variables importantes : l'aire de l'espace étudié et ultimement le nombre



d'espèces distinctes qu'on y trouve. Les images sous-marines sont donc photographiées selon un protocole strict, pour assurer une représentation uniforme dans le temps et dans l'espace du substrat benthique à l'étude (Hodgson *et al.*, 2004). Ensuite, un échantillonnage d'un nombre fixe de zones est extrait de chaque image et est analysé par un expert. Sous chaque zone, la nature de l'objet vivant ou inanimé qu'on y trouve est manuellement identifiée.

Pour les biologistes de l'Université De La Salle, le problème réside principalement dans la quantité d'informations à traiter et en sa complexité très variable. Tout un chacun, sans difficulté, est capable de différencier un amas de sable d'un corail coloré. Toutefois, lorsqu'il s'agit de départager l'*acroporidae acropora hyacinthus* de l'*acroporidae acropora cytherea* de la Figure 0.1, le problème se complexifie énormément. Actuellement, d'importantes ressources humaines spécialisées en biologie sont allouées à cette tâche d'identification qui pourrait être exécutée en grande partie par des acteurs moins spécialisés. En somme, les biologistes devraient plutôt consacrer leurs efforts sur l'analyse du phénomène de l'appauvrissement de la biodiversité marine, ses causes ou éventuellement les pistes de solution pour enrayer le problème que de différencier les spécimens évidents, qui composent la majeure partie des données.

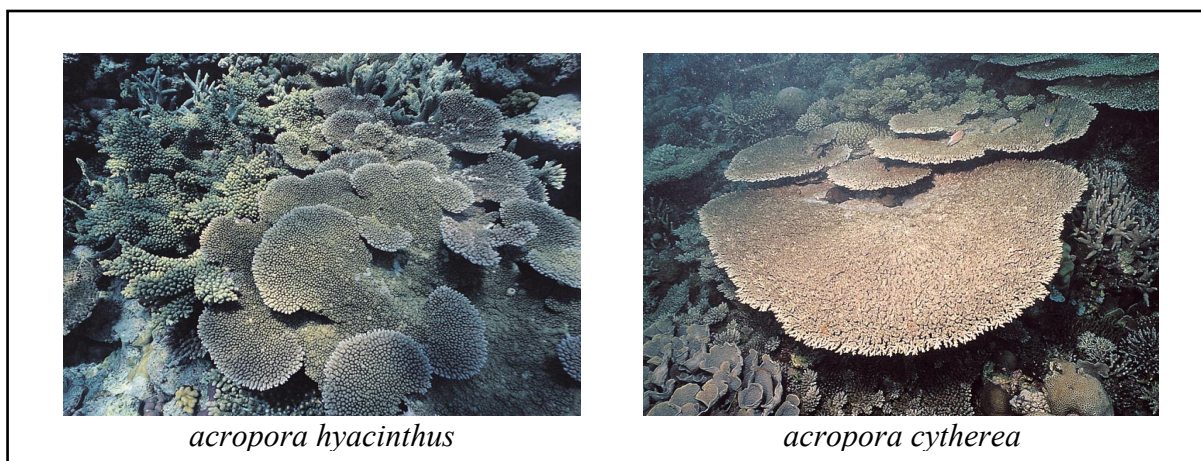


Figure 0.1 Exemple de faible variabilité entre deux espèces.  
Adaptée de Veron et Stafford-Smith (2002)

De plus, l'identification manuelle des spécimens amène tout un lot de problèmes de deuxième ordre. Au cours de ce processus, diverses personnes sont amenées à identifier les images donc, plusieurs méthodes d'identification, ou interprétations de celles-ci, sont utilisées concurremment. Même encadré par un protocole scientifique strict, chaque individu possède ses propres habitudes de travail, son propre niveau d'expérience par rapport à la tâche d'identification et un nombre fini de connaissances par rapport à la taxinomie des coraux. Enfin, une méthode manuelle est sujette à l'erreur humaine, un taux d'erreur très difficile à caractériser statistiquement puisqu'il dépend de tous les facteurs énumérés en plus de la fatigue de l'individu, qui les modifient dans le temps. Dans ce contexte, la crédibilité des données est facilement compromise et cette crédibilité est cruciale pour diriger et justifier les actions qui seront prises pour minimiser l'impact négatif des activités humaines sur la biodiversité.

Toutes ces considérations prises en compte, l'automatisation du processus d'identification des échantillons prélevés est primordiale pour accélérer le traitement des images benthiques, libérer les ressources humaines, formaliser le processus d'identification et assurer la crédibilité des conclusions. Toutefois, même effectuée manuellement, cette tâche d'identification est très complexe à réaliser. Les objets naturels présentent une grande variabilité de formes et de couleurs pour une même espèce et ce phénomène est d'autant plus perceptible chez les coraux. En effet, l'identification des spécimens coralliens est sujette à une énorme variabilité intraclasse (voir Figure 0.2) modulée par de multiples facteurs comme l'âge du spécimen, sa position géographique, les courants marins, la température et la salinité de l'eau, le type d'algue avec laquelle existe une symbiose, la profondeur du benthos, la quantité de lumière disponible et même la période du jour et de l'année où la photographie a été prise. Pourtant pourvu d'une physionomie très simple, le corail montre une étonnante capacité d'adaptation qui donne naissance à de nombreuses configurations différentes pour une même espèce. Ce phénomène paraît peu étonnant lorsqu'on considère le bagage génétique très élaboré des coraux. Selon Miller (2007), ces êtres vivants possèdent un génome plus vaste encore que celui de l'humain.

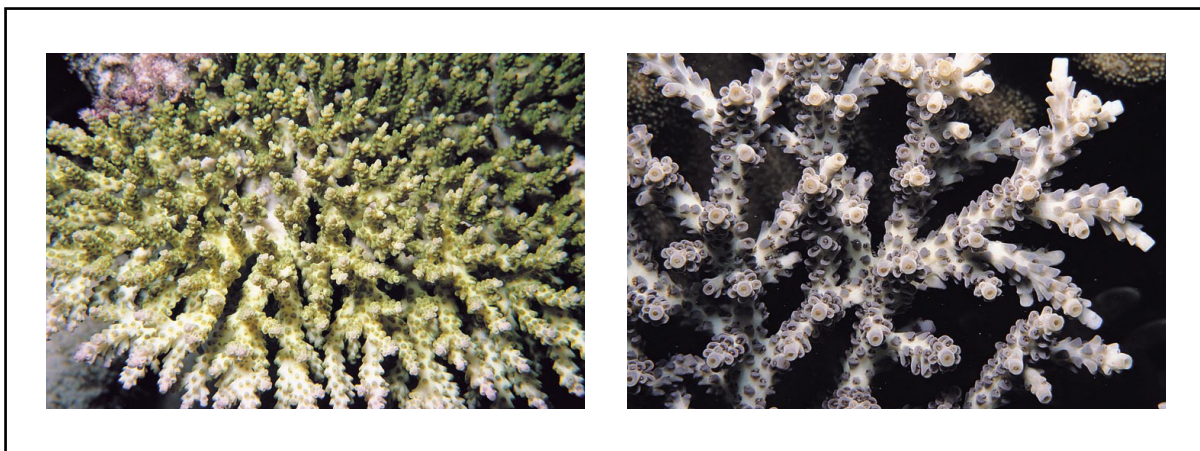


Figure 0.2 Deux spécimens très différents d'*acropora cytherea*.  
Adaptée de Veron et Stafford-Smith (2002)

D'un point de vue de l'automatisation, la tâche est d'autant plus complexe. La plupart du temps, un biologiste chevronné est capable, à partir d'une photo macroscopique couleur, d'identifier correctement l'espèce d'un spécimen corallien. Toutefois, cette méthode est marginalement reconnue par la communauté scientifique. Il n'existe donc aucune documentation sur ce type d'identification visuelle, sur sa méthodologie et sur les caractéristiques discriminantes à observer. Dans la littérature, les informations qui sont utilisées pour déterminer avec précision l'espèce d'un spécimen corallien sont de l'ordre microscopique, et donc non perceptible sur une simple photographie macroscopique. La littérature suggère de discriminer l'espèce par la géométrie du polype, une partie du corail qui est souvent de la taille d'un seul pixel sur les photographies à l'étude (Veron et Stafford-Smith, 2002). Plus encore, dans les cas les plus difficiles, l'identification exacte nécessite une analyse ADN (acide désoxyribonucléique). D'ailleurs, l'avenue des technologies de l'analyse ADN chamboulent la classification classique des coraux. Cette nouvelle source d'informations soulève plusieurs débats dans la communauté scientifique, ce qui complexifie d'avantage la classification puisque la définition des classes évolue constamment (Forsman *et al.*, 2009).

Malgré toutes ces considérations, notre hypothèse de travail stipule qu'*une identification visuelle à l'échelle macroscopique est possible pour départager la majorité des spécimens de*

*corail*. Par conséquent, le présent mémoire fournira une étude des méthodes automatisées pouvant être utiles au suivi de l'intégrité écologique des récifs coralliens. L'objectif global sera donc de fournir un ensemble de méthodes pour solutionner le problème de classification automatisée des coraux à partir de simples photos sous-marines macroscopiques couleurs.

Notre tâche se limite à utiliser des banques d'images de corail déjà existantes, à segmenter ces images automatiquement si nécessaire, à extraire les caractéristiques pertinentes à leur classification et finalement à identifier automatiquement l'espèce des spécimens segmentés à partir de ces caractéristiques. La Figure 0.3 montre ces tâches au sein de la chaîne globale de l'information dans le contexte de notre problématique. De plus, dans le cadre de cette étude nous limiterons notre recherche aux méthodes de vision artificielle (VA) par filtrage spatial et fréquentiel, par analyse morphologique et à la segmentation par analyse de gradients et de régions (Gonzalez et Woods, 2008). Quant au monde de l'intelligence artificielle (IA), nous limiterons la recherche au domaine particulier de la reconnaissance de formes (RDF). Plus particulièrement aux algorithmes de RDF par classification paramétrique, non-paramétrique, par fonctions discriminantes linéaires et aux réseaux neuronaux (Duda, Hart et Stork, 2001).

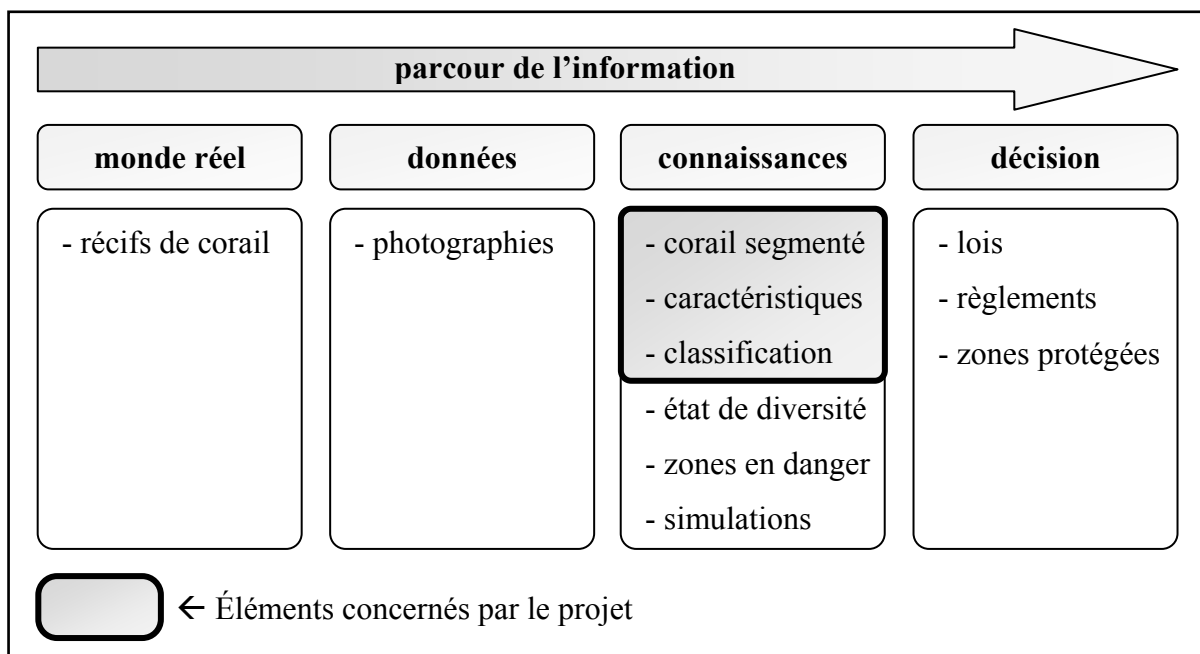


Figure 0.3 Le projet dans son contexte.

Ce travail sera présenté en trois chapitres. En premier lieu, l'état de l'art en reconnaissance automatisée de corail, en RDF et en VA sera présenté. Ensuite, les deux chapitres suivants présenteront la méthodologie et les résultats obtenus considérant l'automatisation de deux approches de la classification biologique des coraux très différentes. Ces chapitres méthodologiques seront suivis d'une conclusion et des recommandations générales traitant de chaque approche. Le Tableau 0.1 présenté ci-dessous détaille l'historique de ce projet qui a débuté tout récemment.

Tableau 0.1 Historique du projet

<b>Date</b>	<b>Événement</b>	<b>Responsable</b>
2007	Collaboration avec l'Université De La Salle	Lael Parrott et Jacques-André Landry
déc. 2007	Projet de fin d'étude au sujet de « [l']optimisation de la taille des régions d'intérêt dans des images non segmentées pour améliorer la performance de classificateurs appliqués aux objets biologiques » <sup>a</sup>	Sébastien Guay
avril 2008	Soutenance de mémoire au sujet de « techniques de l'intelligence artificielle pour la classification d'objets biologiques dans des images bidimensionnelles » <sup>b</sup>	Yan Levasseur
août 2008	Projet de fin d'étude au sujet de « [l']optimisation des méthodes de classification » <sup>c</sup>	Jonathan Bouchard
sept. 2008	Travail sur la description par formes benthiques	Jonathan Bouchard
mars 2009	Rencontre avec M. Wilfredo Y. Licuanan	Jacques-André Landry
juin 2009	Travail sur la description taxinomique complète	Jonathan Bouchard

<sup>a</sup> Guay (2007) <sup>b</sup> Levasseur (2008) <sup>c</sup> Bouchard (2008a)



## **CHAPITRE 1**

### **ÉTAT DE L'ART**

Ce mémoire traite de l'automatisation de la perception visuelle : la vision artificielle (VA); et de l'extraction automatisée de connaissances à partir des éléments perçus : l'intelligence artificielle (IA). Ces deux domaines d'études sont très vastes et font l'objet de plusieurs publications scientifiques. Ce chapitre dressera un portrait des méthodes utiles à notre cause qui sont généralement reconnues dans ces deux champs de connaissances. Les méthodes théoriques présentées seront appliquées au problème de détection et d'identification de spécimens de corail dans des images numériques acquises en plongée sous-marine. La première section de l'état de l'art présentera un résumé de quelques articles scientifiques qui utilisent la cognition artificielle dans le contexte de reconnaissance des coraux. Suite à cette présentation, les diverses méthodes utilisées par l'approche que nous proposons seront détaillées. Les travaux de la littérature nous permettront de comparer nos résultats et d'inspirer la recherche d'une solution à la problématique envisagée.

#### **1.1 Les recherches précédentes sur la reconnaissance automatisée de corail**

Bien que la problématique du suivi écologique des coraux soit d'actualité, il n'existe que très peu de publications scientifiques relatant d'un système capable d'automatiser les tâches liées à ce sujet. Les quelques études qui relatent de l'automatisation des tâches du suivi des communautés benthiques mettent à profit diverses technologies. Les méthodes de la littérature sont distribuées en deux catégories : l'analyse globale, par télédétection; et l'analyse locale, par photographie sous-marine.

Parmi les catégories, divers médias ont été utilisés : l'imagerie multispectrale (voir 1.3.1.3), la vision stéréoscopique (voir 1.3.1.4), l'imagerie par ultra-sons et la photographie traditionnelle. Plusieurs articles relatent même l'utilisation d'une combinaison des médias. Outre la photographie traditionnelle, les médias moins accessibles permettent d'ajouter une information supplémentaire à la prise d'image, soit la réponse à une illumination particulière,

ou encore la topologie tridimensionnelle des spécimens étudiés. Ces informations supplémentaires aident certes à départager les espèces présentes, mais nécessitent de l'équipement hautement spécialisé et coûteux pour effectuer la collecte des données.

*Note : Tout au long de cette section, lorsqu'une terminologie relative à la reconnaissance de formes ou la vision artificielle traitée par ce mémoire est utilisée, une référence à la section correspondante est effectuée.*

### **1.1.1 Les recherches basées sur la télédétection**

Cette catégorie de recherches est incontournable puisque la télédétection (par imagerie satellitaire ou aérienne) est une méthode très populaire actuellement pour effectuer le suivi de la vie aquatique. Les travaux de Lesser et Mobley (2007) présentent un portrait de la littérature au sujet de la classification du substrat marin à l'aide de photographies aériennes hyperspectrale. Cet article présente bien la problématique du suivi écologique des coraux en utilisant ce média. En effet, pour assurer une lecture pertinente de la réponse spectrale, plusieurs facteurs doivent être pris en compte telle la « colonne d'eau » qui sépare l'observateur du sujet. C'est un phénomène connu, l'eau de mer possède un lot de propriétés optiques inhérentes à son milieu et à ses contaminants. Selon ses propriétés, l'eau atténue par absorption certaines longueurs d'onde et cet effet s'accroît avec la taille de la colonne d'eau qui sépare observateur du sujet (Maritorena, Morel et Gentili, 1994). Par extension, la position angulaire de la source d'éclairage fait également varier la réponse spectrale. De ce fait, l'angle de zénith du soleil doit donc également être pris en compte dans les calculs pour corriger la lecture du spectre réfléchi par le substrat benthique.

Pour mener à bien les expériences, les propriétés optiques inhérentes à l'eau de mer de la zone à l'étude sont évaluées en plongée sous-marine avec un imageur multispectral adapté. La bathymétrie (topologie du substrat marin) est acquise à l'aide d'un sonar. Finalement, la nature du benthos est validée en plongée par transects. Lesser et Mobley présentent la relation qui unit la bathymétrie, la nature du benthos et la réponse spectrale observée. Ils



affirment qu'en connaissant deux de ces trois informations, il est possible de déterminer avec une précision acceptable l'information manquante par calcul. Bien que cette méthode soit moins précise que le suivi manuel des communautés benthiques par plongée sous-marine, elle représente un gain significatif puisque l'effort nécessaire à l'acquisition d'images aériennes est de beaucoup inférieur au suivi des communautés benthiques par plongée sous-marine.

Toutefois, il apparaît rapidement que le niveau de détail des approches par télédétection n'est pas suffisant à l'étude de la biodiversité. En effet, les résultats de la classification ne permettent de départager sur l'imagerie aérienne que les coraux, les algues, les plantes aquatiques (macrophytes), le sable et l'aliol (roche sédimentaire). Avec ce niveau de détail, il est possible de déterminer l'abondance et la distribution de la vie aquatique à l'échelle planétaire ainsi que le changement temporel et spatial de ces deux variables. Cependant, cette vue d'ensemble ne permet pas le suivi de la diversité biologique, qui s'intéresse à la répartition locale des espèces. C'est pourquoi nous n'avons pas poursuivi de plus amples recherches dans ce domaine.

### **1.1.2 Les recherches basées sur l'imagerie sous-marine**

Il est certes plus complexe et plus coûteux de transporter une équipe de biologistes et son équipement sophistiqué d'acquisition de données sous l'eau, à proximité des récifs. Toutefois, c'est uniquement en condition de plongée sous-marine que l'identification de l'espèce et le dénombrement des spécimens sont vraisemblablement possibles. Plusieurs publications qui visent l'automatisation du suivi des communautés benthiques relatent cependant l'utilisation d'un VSA (véhicule sous-marin autonome). Dans ce contexte, un petit sous-marin équipé de multiples senseurs est utilisé pour mener à bien la collecte des données automatiquement.

Ce type d'équipement hautement sophistiqué a été utilisé dans les travaux de Johnson-Roberson et al. (2006), où les auteurs ont implanté un algorithme de segmentation et de

classification de spécimens de corail directement au cœur de l'ordinateur de bord du VSA. En mettant à profit la stéréovision, des descripteurs de texture basés sur les ondelettes de Gabor et un SVM (séparateur à vaste marge, voir 1.3.5.5) à titre d'algorithme de reconnaissance de formes (RDF), ils ont réussi à implémenter un système capable de segmenter et d'identifier le benthos en temps réel. La classification présentée par cette recherche ne permet de distinguer que trois formes benthiques distinctes (voir 2.1). Malgré la contrainte de temps réel, ce système admet une performance en classification impressionnante de 95 % suite à l'entraînement en validation croisée (voir 1.3.2.3) du SVM sur près de 8000 images stéréoscopiques.

Johnson-Roberson et al. ont également publié un autre travail sur la reconnaissance automatisée de corail (Johnson-Roberson, Kumar et Williams, 2007). Toutefois, un autre type de capteur a été installé sur le VSA : un sonar. Les données en provenance du sonar ont été utilisées pour analyser la géométrie du substrat océanique. La fusion des données tirées des images couleurs et sonars acquises par le véhicule ont été traitées par le même processus d'extraction de caractéristiques de Gabor. Enfin, le SVM a également été utilisé pour classifier les données parmi deux ou quatre classes. Dans ces conditions, les résultats obtenus sont de 95 % et de 80 % de bonne classification pour les problèmes à deux et quatre classes respectivement.

Malgré l'exceptionnelle complexité technologique des approches précédentes, on dénote chez plusieurs autres auteurs les mêmes éléments au cœur de leur méthodologie. L'utilisation d'une technique de présegmentation des spécimens, des descripteurs basés sur la texture et d'algorithmes de RDF élaborés tel le SVM, est constante pour plusieurs travaux. Le travail de Mehta et al. (2007) est un bon exemple de cette tendance. Cette fois, de simples photographies numériques sont utilisées pour alimenter le classificateur. Plus encore, les valeurs de ton des pixels de petites zones carrées (25 pixels de côté) ont directement été utilisées en entrée du SVM. Malgré la simplicité de ce protocole expérimental, les auteurs clament avoir atteint une performance de l'ordre de 95 % de classification correcte

également. Encore une fois, trois formes benthiques distinctes ont été départagées. Toutefois, seulement cent images ont été utilisées pour entraîner le classificateur.

Les travaux de Marcos (2005) relatent aussi l'utilisation d'un classificateur évolué, le perceptron multicouche (voir 1.3.5.6) et de méthodes d'analyse de la texture : les motifs binaires locaux (voir 1.3.4.3). L'acquisition des images a été effectuée dans ce cas à l'aide d'une caméra vidéo. Les images traitées sont donc des images couleur numériques extraites de la trame vidéo. Encore une fois, trois classes d'objets ont été prises en compte, les coraux vivants, les coraux morts et le sable. Suite à l'entraînement du PM, le taux de classification obtenu en généralisation pour cette approche est de l'ordre de 86 % de classification correcte.

D'autres travaux de la littérature utilisent des méthodes plus simples pour classifier les éléments du substrat benthique. Diaz et Torres (2007) utilisent une procédure qui s'apparente à l'algorithme des plus proches voisins (voir 1.3.5.2) pour déterminer si un segment de l'image appartient à l'une ou l'autre des cinq classes de corail prises en compte par le problème. L'image issue d'une photographie sous marine couleur est d'abord présegmentée automatiquement. Ensuite, la couleur moyenne de chaque segment est utilisée pour classifier les segments. Avenant que la classification en fonction de la couleur soit jugée ambiguë, en fonction d'un seuil de décision ajusté à la main, des caractéristiques de cooccurrence des tons de gris (voir 1.3.4.2) sont extraites pour compléter le portrait. Avec cette approche, l'auteur estime une performance en généralisation de 80 % de bonne classification.

Utilisant le même paradigme de classification, l'article de Soriano et al. (2001) propose un système d'extraction de descripteurs utilisant l'information de couleur et de texture des segments d'images. La couleur moyenne dans un espace normalisé est utilisée conjointement à l'analyse de texture par motifs binaires locaux (voir 1.3.4.3). Les données ont ensuite été classifiées avec la règle du plus proche voisin (voir 1.3.5.2). Pour cette recherche, des images issues de transects vidéo ont été utilisées suite à une segmentation manuelle des coraux. Après une validation des performances, le meilleur classificateur obtenu admet un taux de 48 % de bonne classification sur un problème considérant cinq classes. Les auteurs

mentionnent toutefois un constat intéressant : un taux de classification d'environ 80 % est parfaitement acceptable dans le contexte d'évaluation de la santé des récifs coralliens.

Certains auteurs s'intéressent principalement au problème de segmentation du corail dans des images numériques. C'est le cas de Rivera-Maldonado (2005) qui présente par son mémoire de maîtrise une méthode de segmentation de plusieurs types d'images de corail. Les trois bases de données présentées traitent d'images multispectrales acquises en plongée sous-marine, d'images multispectrales acquises par un VSA et finalement d'images issues d'une télédétection. Dans son mémoire Maldonado utilise la réponse d'une transformée de Hough (utilisée ici comme descripteur de la texture locale) et un algorithme de classification automatique (*clustering*) pour extraire et grouper les zones homogènes de l'image d'un point de vue de la texture. Selon les résultats présentés, l'approche proposée admet une performance constante d'environ 70 % de bonne segmentation, et ce peu importe la base de données considérée lorsque le résultat est comparé à une segmentation manuelle des coraux.

S'éloignant encore davantage des algorithmes de RDF, le travail de Gleason, Reid et Voss (2007) utilise quant à lui un seuillage sur les descripteurs de ses images spectrales pour segmenter le corail du non-corail. Selon ce travail, l'analyse des bandes spectrales de longueurs d'onde de 568 nm (nanomètres) et de 546 nm suffit à segmenter les organismes vivants photosynthétiques. L'analyse de la cooccurrence des tons (voir 1.3.4) sur les images de ces deux bandes spectrale complète le portrait. Un seuil sur les descripteurs issus des images spectrales permet de segmenter le corail avec une précision d'environ 80 %. De plus, ce mécanisme permet également de séparer le corail des algues avec une précision un peu moins élevée.

Enfin, une dernière étude a été consultée sur le sujet. Il s'agit d'une comparaison entre la méthode manuelle généralement utilisée par les biologistes pour effectuer le recensement des coraux d'une certaine région et une méthode automatisée issue de l'informatique biomédicale (Palomar et Mendoza, 2006). Les auteurs ont utilisé une solution commerciale nommée Cellenger® développé par la firme Definens qui permet d'analyser les images de

microscopie optique et qui est généralement utilisée pour l'identification et le dénombrement de cellules vivantes. Cette fois-ci, les performances ne sont pas décrites par un nombre absolu. Les auteurs ont plutôt comparé la stabilité de l'approche informatisée avec celle du système manuel. Il semble toutefois que l'approche implémentée avec Cellenger® ait classifié les coraux de la même façon que les humains 77 % du temps. Enfin, l'analyse de variance effectuée à un niveau de confiance 95 % démontre que l'approche informatisée est statistiquement plus stable que l'humain. Cette étude justifie donc davantage l'effort d'implémenter une solution complètement automatisée pour le recensement des spécimens de corail d'une zone d'intérêt.

### **1.1.3 Discussion sur les recherches précédentes**

Les approches les plus prometteuses de la littérature ont plusieurs points en commun : la photographie sous-marine, l'utilisation d'algorithmes de RDF évolués (SVM, PM), l'utilisation de descripteurs basés sur la texture des coraux et finalement, l'implémentation d'une méthode de segmentation, utilisée préalablement à la classification. Ces quelques similitudes guideront l'approche proposée par notre méthodologie.

Sans faire l'unanimité, une autre information tirée de la littérature est que l'information additionnelle de la topologie tridimensionnelle, acquise par stéréovision ou sonar, et l'information spectrale peuvent aider à discriminer entre les classes de corail. Toutefois, ce constat est mitigé, Gleason, Reid et Voss arrivent à bien classifier les coraux parmi deux classes avec un système très simple en profitant du gain d'information d'une analyse spectrale. D'un autre point de vue, sans aucune information additionnelle, Mehta arrive aussi à bien classifier les coraux parmi trois classes avec un système peu complexe.

Il faut toutefois constater qu'il est très difficile de comparer les approches objectivement : aucune base de données publique n'existe pour échanger sur la problématique de la reconnaissance des coraux. Malgré tout, l'automatisation de cette problématique est très souhaitable, comme le dicte la conclusion de Palomar et Mendoza : l'automatisation de

l'identification des spécimens assure une plus grande stabilité statistique que l'identification manuelle.

## **1.2 Les travaux antérieurs à ce mémoire**

Au cours de cette section, les divers travaux qui ont servi de base à l'expérimentation seront présentés. Ces travaux, effectués par d'autres étudiants dans le contexte précis de la reconnaissance de spécimens coralliens, ont inspiré la structure logicielle qui permet l'extraction des caractéristiques et l'entraînement des classificateurs intelligents. Ces classificateurs permettent ensuite d'effectuer l'identification par formes benthiques des spécimens de la base de données automatiquement.

### **1.2.1 L'intelligence artificielle pour la classification d'objets biologiques**

Le travail de M. Levasseur au sujet de la classification d'objets biologiques est au cœur de l'algorithme d'extraction des caractéristiques de couleurs et de textures qui permet la classification des spécimens. Au cours de ses travaux, M. Levasseur a mis au point un lot de descripteurs qui peuvent être extraits d'une image bidimensionnelle couleur qui sont adaptés à la classification d'objets biologiques. Cet ensemble de descripteurs de texture, de couleur et de morphologie a démontré son efficacité sur plusieurs types d'objets biologiques notamment sur : des céréales (erreur de 0,95%), des nœuds de bois (erreur de 20,37%), des feuilles d'arbres (erreur de 26,67%), des grains de pollen (erreur de 4,15%) et des raisins secs (erreur de 0,22%) (Levasseur, 2008).

Pour l'automatisation de la reconnaissance de spécimens coralliens, une partie des descripteurs décrits dans les travaux de M. Levasseur a été utilisée. Le Tableau 1.1, le Tableau 1.2 et le Tableau 1.3 dressent une liste des descripteurs de morphologie, de texture et de couleur présentés par Levasseur et détaillent leur signification.

Tableau 1.1 Les descripteurs de morphologie  
Adapté de Levasseur (2008)

<b>Description</b>
L'aire du segment
Le centroïde du segment
L'axe majeur du segment
L'axe mineur du segment
L'excentricité du segment (ratio des axes)
L'orientation de l'axe majeur du segment
L'aire convexe du segment (l'aire du plus petit polygone convexe dans lequel entre l'objet)
L'aire remplie (l'aire de la région avec tous ses trous éliminés)
Le nombre d'Euler (le nombre d'objets de la région moins le nombre de trous de ces objets)
Le diamètre équivalent (diamètre du plus grand cercle entrant dans la région)
La solidité (l'aire du segment divisée par l'aire convexe du segment)
L'étendue (le ratio entre l'aire du segment et l'aire de la boîte rectangulaire l'englobant)
Le périmètre
La hauteur
La largeur
La complexité (l'aire divisée par le périmètre élevé au carré)

Les descripteurs de morphologie utilisés sont tous issus de Matlab®, plus précisément de la fonction « regionprops » de la boîte à outils « Image processing toolbox ». La description détaillée des calculs est présentée dans la documentation de Matlab®. Nous nous n'attarderons pas à ces calculs, puisqu'ils sont bien présentés dans la dite documentation.

Notez que l'utilisation des descripteurs de morphologie est limitée aux images segmentées. Ces descripteurs ne seront donc utilisés qu'uniquement dans ce contexte précis.

Tableau 1.2 Les descripteurs de texture  
Adapté de Levasseur (2008)

<b>Description</b>
Moyenne des tons de gris <sup>a</sup>
Écart type des tons de gris <sup>a</sup>
Second moment de l'histogramme normalisé <sup>a</sup>
Coefficient d'asymétrie de l'histogramme des tons de gris <sup>a</sup>
Coefficient d'aplatissement de l'histogramme des tons de gris <sup>a</sup>
Uniformité de l'histogramme des tons de gris <sup>a</sup>
Entropie des tons de gris <sup>a</sup>
Contraste horizontal, vertical, diagonal <sup>a</sup>
Corrélation horizontale, verticale, diagonale <sup>a</sup>
Énergie horizontale <sup>a</sup>
Homogénéité horizontale <sup>a</sup>
Second moment normalisé de la forme binaire locale <sup>a</sup>
Asymétrie de la forme binaire locale <sup>a</sup>
Coefficient d'aplatissement de la forme binaire locale <sup>a</sup>
Origine de la droite pour transformée de Fourier <sup>b</sup>
Pente de la droite pour transformée de Fourier <sup>b</sup>
Corrélation de la droite pour transformée <sup>b</sup>

<sup>a</sup> Voir section 1.3.4; <sup>b</sup> Voir ci-dessous.

Certains des descripteurs du Tableau 1.2 méritent une attention particulière. En effet, les descripteurs basés sur la transformée de Fourier ne sont pas décrits par la littérature. Il s'agit de nouveaux descripteurs proposés par le Dr Jacques-André Landry et présentés au sein du travail de M. Levasseur. Pour calculer les trois dernières caractéristiques : l'origine, la pente et la corrélation d'une droite résumant la transformée de Fourier, Landry propose de faire le cumul de l'énergie fréquentielle à toutes les fréquences disponibles de l'espace de Fourier et d'approximer la fonction cumulative résultante par une droite.



Voici, en détail, comment calculer ce descripteur de texture. Tout d'abord, la transformée de Fourier discrète centrée est extraite de l'image en niveaux de gris, tel que présenté à la section 1.3.4.4. Ensuite, la sommation de l'énergie fréquentielle de chaque bande est effectuée. Chacune des bandes de fréquence d'une transformée de Fourier est distribuée sur un cercle centré sur l'image transformée. Plus le rayon de ce cercle est grand, plus la fréquence étudiée est élevée. Enfin, l'allure de la courbe cumulative de l'énergie en fonction de la fréquence est approximée par une droite par régression linéaire, dont les paramètres sont conservés à titre de descripteurs numériques de la texture.

Tableau 1.3 Les descripteurs de couleur  
Adapté de Levasseur (2008)

<b>Description</b>
Le mode de l'histogramme du canal couleur (sur R, G, B) <sup>a</sup>
Valeur moyenne de l'intensité du canal couleur (sur R, G, B) <sup>a</sup>
Valeur maximale de l'intensité du canal couleur (sur R, G, B) <sup>a</sup>
Valeur minimale de l'intensité du canal couleur (sur R, G, B) <sup>a</sup>
Somme du résultat d'un filtre Sobel du canal couleur (sur R, G, B) <sup>b</sup>
Entropie du canal couleur (sur R, G, B) <sup>a</sup>

<sup>a</sup> Voir section 1.3.4, <sup>b</sup> Voir section 1.3.3.1.

Pour ce qui est de la couleur, les descripteurs choisis sont de simples statistiques de l'histogramme. Le descripteur utilisant le détecteur de contours de Sobel étudie quant à lui la présence de contrastes francs dans les canaux de chaque couleur. La somme des pixels répondant au critère de Sobel indique si le canal couleur est contrasté (forte somme) ou non (faible somme). Ces statistiques et la méthode de Sobel sont décrites en 1.3.3.1 et 1.3.4.

En plus de fournir l'infrastructure logicielle pour faire l'extraction de ces descripteurs d'images au sein de MATLAB®, le travail de M. Levasseur permet la sauvegarde des données d'une expérience basique au format *attribute-relation file format* (ARFF). Le format ARFF

est typiquement utilisé dans l'environnement Weka, une boîte à outils qui implémente plusieurs algorithmes d'IA (Hall *et al.*, 2009; Witten et Frank, 2005b).

### 1.2.2 L'optimisation de la taille des régions d'intérêt

Au cours de son projet de fin d'études en génie de la production automatisée, M. Sébastien Guay a poursuivi le travail sur l'extraction de caractéristiques pour l'identification automatisée des spécimens de corail (Guay, 2007). Ce travail est basé sur la banque d'images obtenue lors d'une collaboration avec l'Université De La Salle aux Philippines (voir le CHAPITRE 2 pour de plus amples renseignements). Notez que cette base de données contient un lot d'images où dix points d'échantillonnage ont été identifiés manuellement par la forme benthique qui s'y trouve (voir les sections 2.1 à 2.3 pour plus d'informations). Dans ce contexte, le travail de M. Guay avait pour but de déterminer la taille idéale de la zone image à considérer sous chaque marqueur d'échantillonnage pour maximiser les performances en classification à partir des descripteurs de M. Levasseur présentés en 1.2.1. Enfin, c'est au cours de ce mandat que l'infrastructure logicielle pour extraire les zones échantillonnées pour chacun des dix marqueurs par image (initialement développée par Landry) a été formalisée et documentée.

Cette architecture comprend l'utilisation de MATLAB® pour extraire les zones échantillonnées centrées sur les dix points d'échantillonnages (d'une taille variant de 48x36 à 432x324 pixels). L'algorithme permet également de classer les diverses zones images dans une arborescence identifiant le code benthique correspondant. Ensuite, l'extraction de caractéristiques présentée en 1.2.1 entre en action. Chaque valeur descriptive est stockée sur une ligne du fichier de sortie au format ARFF et le tout est accompagné du code benthique de l'échantillon. Enfin, entre en jeu divers algorithmes de RDF implémentés par Weka (Hall *et al.*, 2009; Witten et Frank, 2005b). Dans le cadre de cette expérience, le séparateur à vaste marge (SVM) et l'algorithme des plus proches voisins (kPPV) ont été utilisés (voir section 1.3.5). Aucune information n'est disponible quant aux métaparamètres utilisés.

L'étude conclut en affirmant qu'aucune des tailles concevables pour la zone d'échantillonnage n'est significativement optimale. Bien qu'une légère amélioration des performances soit observée lorsque l'aire de zone échantillonnée s'accroît, le meilleur résultat obtenu n'est toujours pas suffisant pour établir un système adéquat d'identification automatisé. Le meilleur système obtenu commet une erreur de classification sur près de 60 % des spécimens testés. Cette conclusion ouvre la porte à plus d'investigations sur la base de données et sur le système de classification, qui sont source du problème de performance.

### 1.3 Les sujets abordés en vision artificielle et reconnaissance de formes

Pour bien décrire les sujets reliés à l'élaboration d'un système de RDF, les diverses sections de cette revue de littérature seront structurées selon les six étapes fondamentales du traitement des données. La Figure 1.1 présente l'architecture d'un système de reconnaissance de forme utilisé sur les données issues d'une procédure de VA. Elle est sans aucun doute la plus importante de ce travail puisque tant la structure logicielle utilisée que la structure même de ce document s'inspirent de cette architecture.

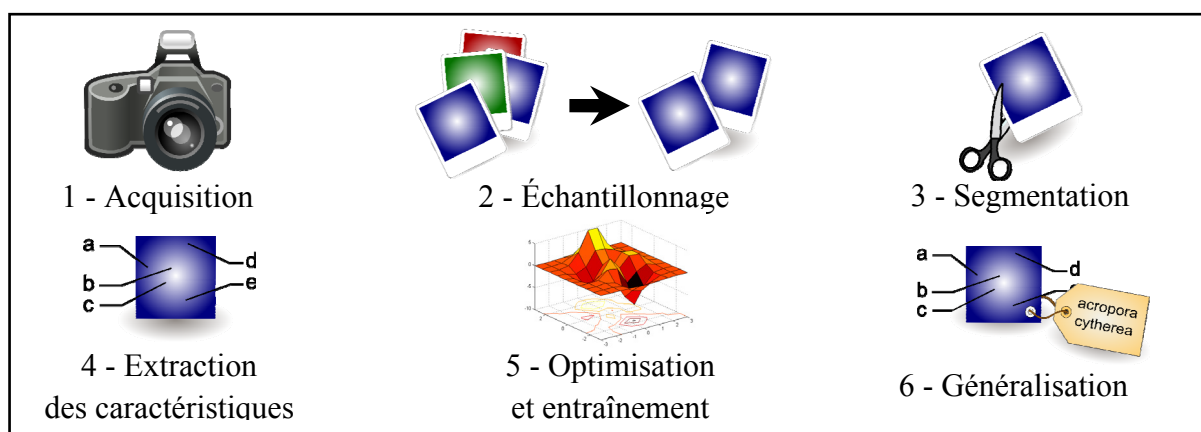


Figure 1.1 Schéma d'un système de reconnaissance de formes par vision artificielle.

### 1.3.1 L'acquisition des images et les prétraitements

Avant toute étude scientifique, une collecte d'information est nécessaire. Dans le domaine de la VA, cette collecte d'information prend la forme d'une séance photo. Les équipements utilisés à cette étape impliquent déjà beaucoup de concepts techniques d'importances qui ont un impact sur les résultats obtenus. En photographie numérique, en plus de la focale et de l'ouverture de la lentille, plusieurs éléments de l'image sont souvent ajustés automatiquement à chaque prise de vue, telles la balance des couleurs (parfois nommée balance des blancs) et la plage dynamique de l'image. Cette section traitera des principaux éléments techniques liés à cette étape primordiale.

#### 1.3.1.1 La balance des couleurs

La balance des couleurs est un traitement qui uniformise la teinte des objets, peu importe la source d'éclairage utilisé. Il existe de nombreux algorithmes automatiques qui tentent de détecter la nature de la source d'éclairage et de corriger son effet, tel celui de Liu, Chan et Chen (2002). Toutefois, dans un contexte plus formel d'utilisation des images, la balance des blancs manuelle est préférable, puisqu'elle est plus stable. Pour calibrer manuellement l'appareil photographique, l'utilisateur doit photographier une référence blanche. Après l'acquisition de l'image, l'appareil photo pondère la réponse RVB (rouge, verte et bleue) de chacun de ses capteurs pour assurer la neutralité du spécimen blanc photographié, ce qui équilibre les couleurs pour les prises de vues subséquentes. L'algorithme généralement utilisé pour accomplir cette tâche s'exécute en deux étapes: - 1 - la segmentation de la zone blanche de l'image et - 2 - le calcul de la pondération des canaux RVB de l'image qui permet de neutraliser les tons pour la zone segmentée.

La pondération est appliquée à chaque canal couleur à l'aide d'une transformée linéaire simpliste telle qu'illustrée à l'équation (1.1) tirée de Wikipedia (Wikimedia Foundation, 2010). Dans cette équation, les canaux couleurs de l'image originale (avant correction des couleurs) sont notés par  $R'V'B'$  et ceux de l'image corrigée par  $RVB$ . La valeur moyenne de

ton de la zone considérée blanche préalablement segmentée est notée par  $R'_b V'_b B'_b$ . Notez que la valeur 255 ( $2^8 - 1$ ) correspond à la valeur maximale que peut prendre un canal couleur dans une image codée sur huit bits.

$$\begin{bmatrix} R \\ V \\ B \end{bmatrix} = \begin{bmatrix} 255/R'_b & 0 & 0 \\ 0 & 255/V'_b & 0 \\ 0 & 0 & 255/B'_b \end{bmatrix} \begin{bmatrix} R' \\ V' \\ B' \end{bmatrix} \quad (1.1)$$

Suite au traitement, les couleurs représentées par les images sont fidèles à la réalité. Ce traitement est donc primordial lorsque l'information véhiculée par la couleur doit être utilisée par la suite pour prendre une décision sur la nature des objets qu'on y trouve.

### 1.3.1.2 L'ajustement de la plage dynamique

Pour assurer des contrastes vifs et réalistes dans les images acquises, les appareils photo numériques utilisent plusieurs méthodes de correction de l'illumination. En plus des paramètres habituels tels l'ouverture de la lentille et le temps d'exposition, la sensibilité du capteur peut être ajustée automatiquement pour centrer les diverses intensités lumineuses perçues dans la plage de réponse du capteur. La sensibilité ISO (*International Organisation for Standardisation*) ou vitesse ISO en photographie est un gain appliqué à l'information perçue par le capteur image qui ajuste le compromis entre la sensibilité du capteur en faible lumière et le bruit occasionné par cette sensibilité accrue. En plus du gain, certains appareils contrôlent également l'histogramme d'intensité lumineuse en rééchantillonnant les données perçues pour qu'elles occupent une plus grande part de la plage dynamique permise par la profondeur de couleur choisie (souvent représentée à l'interne par un entier de huit bits).

Ces deux traitements affectent l'histogramme des tons de l'image. Un étirement d'histogramme est l'effet observé et se solde en une augmentation des contrastes sur l'image corrigée. Pour étirer l'histogramme des tons sur toute la plage dynamique disponible, une

simple transformée linéaire est encore une fois la solution. L'équation (1.2) présente ce traitement.

$$\begin{bmatrix} R \\ V \\ B \end{bmatrix} = \frac{255}{(m_x - m_n)} \begin{bmatrix} R' - m_n \\ V' - m_n \\ B' - m_n \end{bmatrix} \quad (1.2)$$

Encore une fois ici, les canaux couleurs de l'image originale (avant correction des couleurs) sont notés par  $R'V'B'$  et ceux de l'image corrigée par  $RVB$ . La valeur de ton minimale, tout canal couleur confondu est notée  $m_n$  et la valeur maximale  $m_x$ . Également, la valeur 255 ( $2^8 - 1$ ) correspond à la valeur maximale que peut prendre un canal couleur dans une image codée sur vingt-quatre bits (avec trois canaux couleurs codés sur huit bits).

### 1.3.1.3 La spectrométrie

En plus des concepts liés à la photographie numérique classique, plusieurs autres technologies peuvent être mises à profit pour tirer davantage d'informations d'une image. L'imagerie multispectrale est une de ces technologies qui permet d'analyser la réponse particulière d'un objet à une certaine longueur d'onde, qui peut être comprise dans le spectre du visible, ou même hors du spectre visible à l'œil humain. L'information spectrale est importante dans la problématique de l'identification du corail. Les coraux, comme beaucoup d'espèces vivantes, dépendent en partie de l'énergie du soleil pour vivre. De ce fait, chaque espèce absorbe certaines longueurs d'onde et en réfléchit d'autres. Cette signature spectrométrique est discriminante quant à l'identité taxinomique d'un spécimen. Comme mentionné en 1.1.2, les travaux de Gleason, Reid et Voss (2007) indiquent qu'une certaine plage du spectre visible peut être utilisée pour départager les coraux des algues et du substrat océanique.

Toutefois, l'analyse du spectre réfléchi nécessite des outils sophistiqués et d'autant plus dispendieux, lorsque l'on considère que ces outils doivent être utilisables sous l'eau. En plus de résister à ces conditions physiques, les outils doivent tenir compte de l'effet de l'eau sur le

spectre lumineux de l'éclairage de la scène. L'eau absorbe une partie du spectre visible et cet effet s'accroît avec la distance qui sépare l'observateur du sujet.

#### **1.3.1.4 La stéréovision**

Une autre technologie d'avant-garde peut augmenter significativement l'information contenue par l'image : la vision stéréoscopique. Cette technologie permet de capter une image numérique qui véhicule l'information tridimensionnelle. Idem à la vision humaine, la stéréovision est implémentée à l'aide de deux caméras séparées par une distance fixe et permet donc de modéliser partiellement la scène dans un espace à trois dimensions via divers calculs trigonométriques. Une représentation complète peut être effectuée si plusieurs plans de caméras sont utilisés pour couvrir le sujet de toutes parts. L'information tridimensionnelle est également un atout dans la reconnaissance de spécimens de corail, puisque la forme de croissance est une des caractéristiques discriminantes de l'identité taxinomique.

#### **1.3.2 L'échantillonnage**

Pour assurer la précision des estimés statistiques sur une population donnée, toute analyse scientifique se doit de considérer la bonne méthode d'échantillonnage. Dans le contexte de cette étude, cette étape primordiale peut être décrite selon deux points de vue différents : celui des biologistes et celui de notre problématique d'automatisation de l'identification des spécimens. Dans le contexte biologique, l'évaluation de la biodiversité de la population des coraux nécessite un protocole strict d'échantillonnage sur le terrain. Cet échantillonnage est hors de notre mandat, toutefois pour de plus amples informations les guides de Page et al. (2001) et de Hodgson et al. (2004) pourront vous renseigner sur le sujet. Du point de vue de l'automatisation des tâches d'identification par VA et IA, divers thèmes importants doivent être maîtrisés pour bien conclure sur la viabilité d'une approche automatisée. Les prochaines sections décriront les divers outils statistiques utilisés pour valider un système de RDF.

### **1.3.2.1 La validation simple**

Dans un premier temps, afin de vérifier le bon fonctionnement d'une approche, les données disponibles doivent être séparées en plusieurs segments indépendants. L'algorithme de reconnaissance de formes est d'abord entraîné sur une partie des données (typiquement les deux tiers des données disponibles) et dans un deuxième temps, l'évaluation de la performance du système est conduite sur la portion restante. Ainsi, il est possible de vérifier si le système est capable de traiter correctement des données qui n'ont jamais été vues auparavant. La faculté d'un système à bien gérer de nouvelles données est nommée la capacité de généralisation. En résumé, la validation simple met en jeu deux segments de données : la base de données d'apprentissage et la base de test (Witten et Frank, 2005a).

Parfois, il peut être nécessaire de réserver un troisième segment de données pour accomplir la validation des métaparamètres. Cette procédure permet d'optimiser certains paramètres de l'expérience de façon indépendante de l'apprentissage et des tests.

### **1.3.2.2 La stratification**

Au cours d'une procédure de validation, il peut arriver par malchance que le segment utilisé pour l'entraînement (ou les tests) soit non représentatif de la base de données complète. Par exemple, il pourrait arriver qu'une des classes du problème ne soit pas représentée dans l'un ou l'autre des segments de données. Un tel problème de représentation des classes peut nuire grandement à la capacité de généralisation de l'algorithme, ou encore pire, fausser la conclusion tirée des résultats de test. Pour éviter ce problème, une procédure de stratification permet de sélectionner un nombre approximativement proportionnel d'échantillons de chaque classe pour chaque segment de données. Ainsi, les proportions entre les classes pour chaque segment de données sont similaires à celle de la base de données entière (Witten et Frank, 2005a).



### 1.3.2.3 La validation croisée

Enfin, la validation simple avec stratification n'est généralement pas suffisante pour éliminer tout biais dans les résultats obtenus. En fait, pour assurer la représentativité des performances observées et par le fait même de la capacité de généralisation du système, il faut répéter maintes fois l'expérimentation avec un échantillonnage différent. Un nombre suffisant de répliques fondées sur autant d'échantillonnages aléatoires stratifiés permet d'atténuer le biais lié au rééchantillonnage des données.

La validation croisée est une procédure qui structure les itérations nécessaires à l'évaluation des performances du système. Avant toute chose, il faut déterminer le nombre de répliques désirées. Typiquement, la validation croisée est effectuée sur dix itérations, on parle alors d'une validation croisée à dix partitions. Dans un premier temps, la base de données doit être séparée en dix partitions indépendantes stratifiées d'une taille approximativement égale. À la première itération, les neuf premières partitions sont utilisées pour l'entraînement du classificateur, et la dernière est utilisée pour les tests. Ensuite, les partitions 2 à 10 sont utilisées pour l'entraînement et la partition 1 est utilisée pour les tests. À la troisième itération, les partitions 3 à 10 et la partition 1 sont utilisées pour l'entraînement, et la partition 2 est utilisée pour les tests...

L'algorithme entraîne ainsi dix fois l'algorithme de RDF sur autant de segments de données différents et dix tests indépendants sont conduits par la suite. Le nombre de partitions (itérations) peut être modifié pour répondre aux besoins particuliers d'une expérimentation. Witten et Frank (2005a) mentionnent qu'une validation à 5 ou 20 partitions est probablement un choix aussi valide qu'un nombre fixé à 10.

### 1.3.3 La segmentation

À cette étape, nous avons à notre disposition une banque d'images et un échantillonnage de ces images qui permet une validation statistique d'un système de RDF. L'étape subséquente

est de localiser dans chaque image, le sujet d'intérêt (si nécessaire). Ce problème est typiquement du domaine de la VA. Toutefois, pour automatiser complètement la tâche, divers outils de la RDF peuvent être mis à contribution. Cette section traitera des outils de segmentation automatisée de la littérature qui peuvent être mis à contribution pour résoudre la segmentation des coraux dans des images numériques acquises en milieu naturel.

### 1.3.3.1 Les fondements de la vision artificielle

Avant d'entamer une description des méthodes de segmentation, quelques notions de base doivent être considérées. Tout d'abord, comme illustré par les équations de la section 1.3.1, le traitement d'images par VA considère une image couleur comme une matrice à trois dimensions. Les deux premières dimensions sont nommées  $x$  et  $y$ , soit la largeur et la hauteur de l'image en pixels. La dernière dimension, nommée  $c$ , correspond au nombre de couches spectrales. Dans une image RVB, la dimension  $c$  est de cardinalité 3. Chacune des trois couches  $c$  représente respectivement l'intensité du rouge, du vert et du bleu dans cet espace de couleur additif (où l'on obtient le blanc pur lorsque chaque intensité de couleur est à son maximum). Avec le paradigme matriciel, diverses opérations mathématiques peuvent être conduites pour manipuler les données.

La corrélation (souvent, mais faussement appelée *convolution*) matricielle est une des opérations les plus importantes à maîtriser. La corrélation permet d'effectuer d'innombrables opérations locales sur l'image. Pour effectuer ce traitement, un masque de corrélation est d'abord défini. Le masque est une matrice plus petite que l'image originale qui est déplacée de façon à ce que son centre soit juxtaposé successivement avec tous les pixels de l'image à traiter. Dans le contexte des opérations de filtrage linéaire, à chaque déplacement, la somme des valeurs de tons de l'image originale, pondérées par les coefficients du masque, est consignée dans l'image de sortie (voir la Figure 1.2, pour un exemple concret). Les valeurs obtenues sont ensuite arrondies à l'entier le plus près. Notez qu'un traitement particulier est nécessaire pour les bordures de l'image. Souvent, la valeur des bordures est simplement répliquée de part et d'autre pour permettre au masque de traiter l'image en entier.

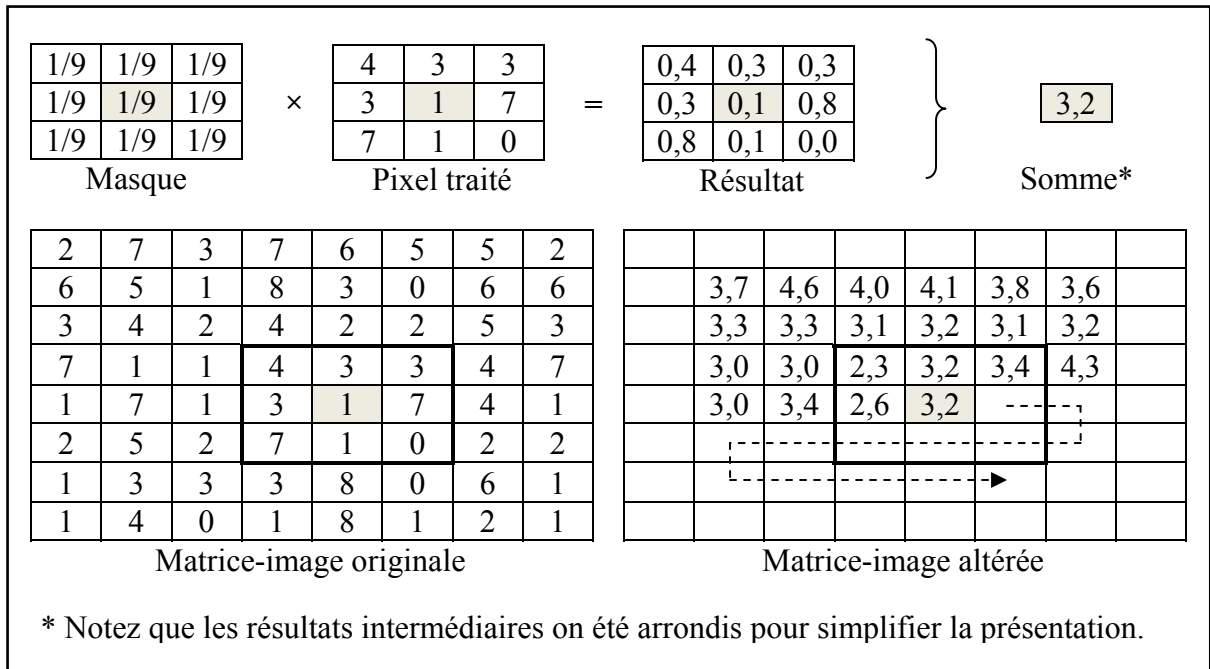


Figure 1.2 Corrélation d’une image de 3 bits par un masque moyennant de taille 3.

En fonction des coefficients et de la taille du masque de filtrage, de multiples opérations peuvent être accomplies. Voici un court résumé des opérations de filtrage couramment utilisées en VA. Le Tableau 1.4, inspiré de Gonzalez et Woods (2008) et de (Landry et Bouchard, 2007) ci-dessous présente ces opérations.

Tableau 1.4 Quelques masques de corrélation pour le filtrage

Nom du filtre	Masque	Description
Moyennant	$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	Adoucit et réduit les contrastes.
Sobel	$\begin{bmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$ ou $\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$	Met en évidence les discontinuités horizontales ou verticales.
Laplacien	$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	Approximation d’un laplacien, dérive l’image, repère les discontinuités.
Max ou min	$\max(v)$ ou $\min(v)$	Élimination du bruit poivre ou sel.

Pour d'autres types de traitements, une image numérique est souvent considérée comme un ensemble de tons qui répondent à une fonction de densité de probabilité. Ainsi divers outils statistiques peuvent être utilisés pour manipuler les données et en extraire l'information pertinente. Dans ce paradigme, l'image est modélisée par un histogramme à  $L$  bandes. Le nombre de bandes  $L$  est déterminé par la profondeur de couleur. Par exemple, dans une image RVB standard d'une profondeur de couleur de 24 bits, soit 8 bits par canal couleur,  $L$  prends la valeur  $2^8$  et donc 256 bandes sont échantillonnées sur l'histogramme de chaque canal.

L'étirement d'histogramme tel que présenté par l'équation (1.2) est un exemple des opérations qui peuvent être effectuées grâce à cette technique. Divers autres traitements et divers descripteurs peuvent être tirés des statistiques de l'histogramme (ou de la fonction de densité) des tons de gris. Les prochaines sections de ce document traiteront de certains d'entre eux.

### 1.3.3.2 Le seuillage

Maintenant que quelques concepts fondamentaux sont définis, voici la présentation d'une première méthode de segmentation. La méthode de segmentation la plus simple est sans aucun doute le seuillage de l'histogramme. Comme le nom de la méthode l'indique, il suffit de sélectionner une valeur seuil pour séparer en deux classes les pixels d'une image numérique. La Figure 1.1 démontre un exemple classique de ce type de traitement tel qu'il est présenté dans la documentation de MATLAB®. Pour cette opération, une valeur de seuil de 126 a été sélectionnée, ce qui permet de séparer les pièces de monnaie de la trame de fond de l'image. Tous les pixels dont l'intensité est supérieure à 126 se voient attribuer la classe 1, et les autres 0. Ainsi, l'image binaire résultante illustre les pièces en blanc et l'arrière-plan en noir.

Ce type de seuillage peut être automatisé et ainsi, la valeur de seuil qui minimise la variance intraclasse (et par le fait même, maximise la variance interclasses) est obtenue

automatiquement. La méthode proposée par Otsu (1979) est basée sur ce critère et s'exécute comme illustrée à la Figure 1.4.

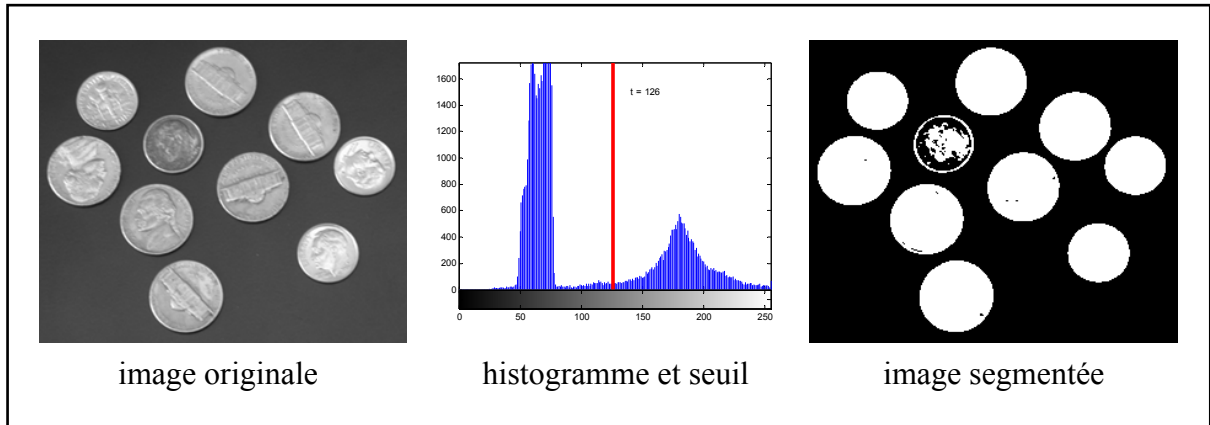


Figure 1.3 Exemple de seuillage.  
Image originale tirée de MATLAB®

**Variables :**  $\omega_i \rightarrow$  probabilité de la classe  $i$ ,  $\mu_i \rightarrow$  moyenne de la classe  $i$ ,  
 $\sigma_b^2 \rightarrow$  variance inter-classe

1. Calculer l'histogramme et les probabilités de chaque niveau d'intensité
2. Initialiser  $\omega_i(0)$  et  $\mu_i(0)$
3. Pour chaque seuil possible  $t \in \{0, 1, \dots, 255\}$ 
  - a. Mettre à jour  $\omega_i(t)$  et  $\mu_i(t)$
  - b. Calculer  $\sigma_b^2(t) = \omega_1(t)\omega_2(t)[\mu_1(t) - \mu_2(t)]^2$
4. Trouver le seuil idéal qui maximise la variance inter-classe :  
 $t_i | \max(\sigma_b^2) = \sigma_b^2(t_i)$

Figure 1.4 L'algorithme d'Otsu.  
Adaptée de Wikipedia (Wikimedia Foundation, 2010)

Notez que c'est le seuil optimal d'Otsu qui a été sélectionné pour effectuer le seuillage dans l'exemple de la Figure 1.3. La valeur 126 (relative au maximum possible de 255) est donc le seuil qui minimise la variance inter-classe.

### 1.3.3.3 Le laplacien du gaussien

Le LoG (laplacien du gaussien) est une méthode de détection de discontinuités classique de la littérature. Cet opérateur est souvent utilisé pour faire la détection de taches (*blobs*) dans les images numériques puisqu'il est possible de paramétrer sa sensibilité. Tout comme le laplacien simple, le LoG s'utilise en notation matricielle avec la procédure de corrélation présentée en 1.3.3.1. Toutefois, pour de meilleurs résultats aucune approximation n'est généralement effectuée avant le traitement. Le masque de filtrage LoG de taille  $n_1$  par  $n_2$  et de variance  $\sigma^2$  est obtenu à partir des équations (1.3) et (1.4) où  $h_g$  est le filtre gaussien et  $h$  le laplacien du filtre gaussien (Gonzalez et Woods, 2008).

$$h_g(n_1, n_2) = e^{-(n_1^2 + n_2^2)/(2\sigma^2)} \quad (1.3)$$

$$h(n_1, n_2) = \frac{(n_1^2 + n_2^2 - 2\sigma^2)h_g(n_1, n_2)}{2\pi\sigma^6 \sum_{n_1} \sum_{n_2} h_g} \quad (1.4)$$

Après avoir calculé le filtre de taille et de variance désirée, il suffit de l'appliquer à l'image désirée. La Figure 1.5 montre le résultat de l'application d'un tel filtre sur l'image des pièces de monnaie issue de MATLAB®. Un seuillage simple dans l'image résultante permettrait par exemple d'extraire facilement les contours de chacune des pièces.

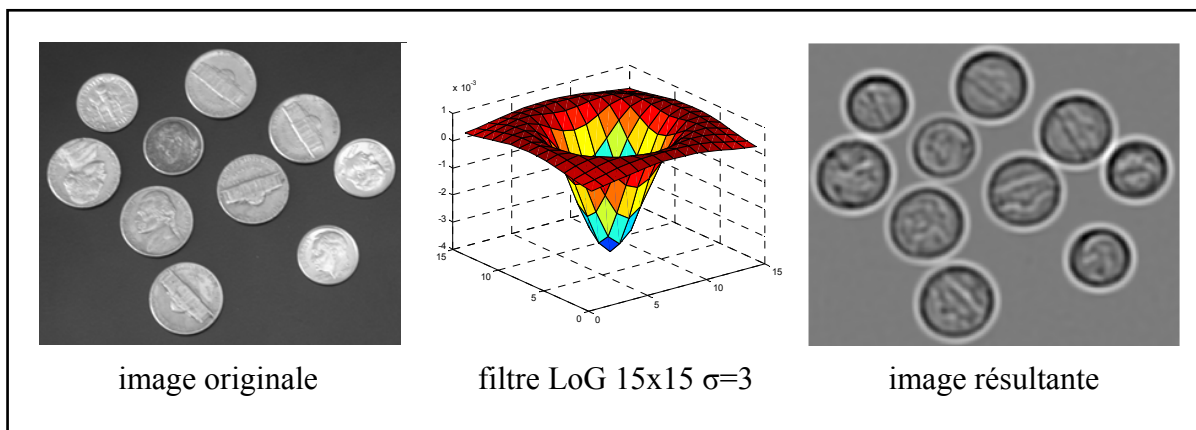


Figure 1.5 Exemple de filtrage LoG.  
Image originale tirée de MATLAB®

### 1.3.3.4 Le partage des eaux

Le partage des eaux (Meyer, 1994) est une autre méthode de segmentation automatisée qui ne dépend cette fois aucunement de notions matricielles ou de statistiques sur l'histogramme. Il s'agit d'un algorithme basé sur la morphologie, il visualise les tons de gris d'une image comme l'altitude topologique d'une surface. Cette surface est ensuite inondée virtuellement jusqu'à ce que les bassins formés par le liquide se fusionnent. Chaque fois qu'un bassin fusionne avec un bassin voisin, une frontière est créée et consignée dans l'image de sortie. Suite à l'exécution de l'algorithme, l'image est segmentée en de multiples zones contigües. Ces zones sont délimitées par un segment blanc dans l'exemple de la Figure 1.6.

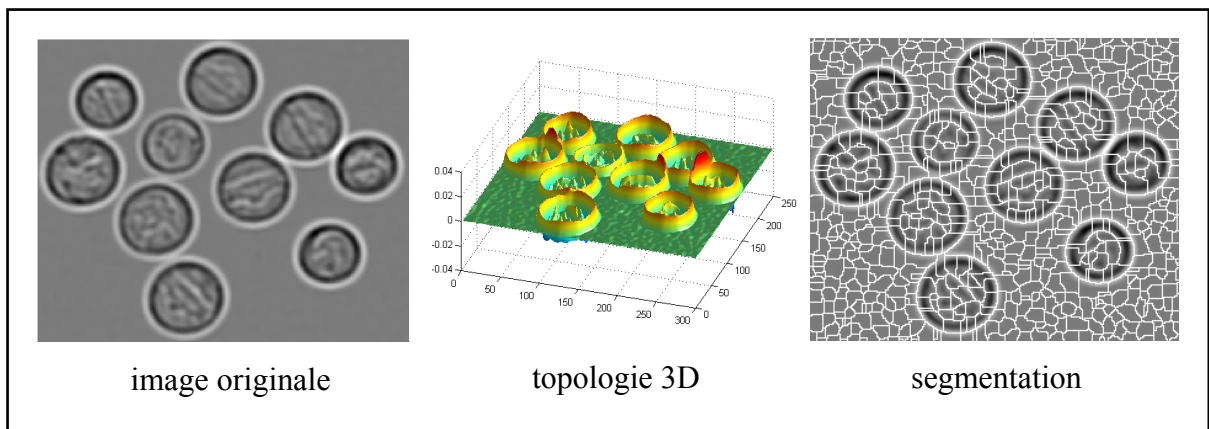


Figure 1.6 Exemple de segmentation par partage des eaux.  
Image originale tirée de MATLAB®

L'algorithme du partage des eaux s'exécute comme suit (Gonzalez et Woods, 2008; Wikimedia Foundation, 2010) :

1. Un lot de marqueurs, à l'endroit où l'inondation doit débiter, est sélectionné. Un libellé numérique est assigné à chacun de ces marqueurs;
2. Les pixels voisins de chaque région marquée sont insérés dans une file de priorité avec un niveau de propriété correspondant à l'intensité des pixels;

3. Le pixel avec le plus haut degré de priorité est extrait de la file. Si les pixels voisins de ce pixel ont tous le même libellé, alors le pixel courant se voit attribuer ce libellé. Tous les pixels voisins non libellés qui ne sont pas déjà dans la file de priorité y sont ajoutés;
4. Répéter l'étape 3, jusqu'à ce que la file de priorité soit vide.

Suite à l'exécution de l'algorithme, tous les pixels qui ne sont pas libellés sont les frontières de segmentation.

### 1.3.3.5 Les méthodes avancées

En plus des méthodes simples, la littérature scientifique contemporaine ajoute un lot de méthodes avancées qui permettent de segmenter des objets complexes dans des conditions particulières. Un auteur et son approche particulière ont retenu notre attention au cours de la revue de littérature. La prochaine section présentera les résumés de ces quelques techniques d'avant-garde tirés de Bouchard (2009a).

La première approche présentée a été mise en place par l'équipe de Todorovic et Ahuja (2009). Cet article propose un algorithme de segmentation capable d'extraire les régions texturées d'une image arbitraire. La méthode proposée utilise une série de techniques statistiques pour modéliser la variance intrinsèque aux textures naturelles et pour caractériser les régions extraites. L'auteur définit la segmentation comme étant la « délimitation des frontières de toutes les textures dans une image arbitraire. »

L'algorithme proposé exige toutefois qu'une condition soit respectée à priori : la granularité de la texture doit être grossière pour qu'un algorithme puisse la segmenter. En d'autres termes, les textures de l'image doivent être constituées d'éléments de textures (de *texels*) qui sont d'une dimension supérieure au pixel.

La principale difficulté liée à la segmentation de textures naturelles réside dans la variabilité des objets. En effet, les *texels* d'une image naturelle ne sont pas identiques, leur répétition



spatiale n'est pas strictement périodique. Considérant cette problématique, l'auteur suggère de caractériser les textures par une fonction de densité de probabilité décrivant la variation naturelle des propriétés géométriques, photométriques, structurelles, d'orientation et de déplacement des *texels*.

L'algorithme proposé se décompose en deux phases importantes. Dans un premier temps, l'image est segmentée avec un algorithme simple. Cette segmentation a pour but d'isoler les différents *texels* de l'image qui seront par la suite attribués à une zone de texture plus large. À même la phase de segmentation, une multitude de descripteurs statistiques sont calculés pour chacune des régions isolées. Par la suite, un algorithme d'agrégation se charge de trouver les groupements dans l'espace des descripteurs. Cet algorithme tente de trouver les zones de l'espace des descripteurs avec la plus forte densité, ces zones représentent un lot de *texels* qui possèdent des caractéristiques similaires. Une fois regroupés, ces ensembles de *texels* semblables forment une texture. Voici en détail l'algorithme de segmentation :

La première étape se charge de la segmentation multirésolutions de l'image et du calcul des descripteurs pour chaque région segmentée (couleur, aire, forme, position). La segmentation préalable est effectuée avec un algorithme spécialement conçu. Cet algorithme permet d'explorer les divers groupements de pixels qui possèdent des caractéristiques visuelles semblables, et ce, à plusieurs échelles. Pour ce faire, le contraste de l'image à traiter est modifié à chaque itération et une nouvelle segmentation est exécutée. L'image est d'abord segmentée très finement et par la suite à divers niveaux jusqu'à l'atteinte d'une segmentation très grossière (où les segments couvrent près de 50 % de l'image). Cette méthode instaure une hiérarchie entre les divers niveaux de segmentation. En d'autres termes, les petites régions extraites au début de l'exécution sont comprises dans les zones plus grandes extraites lors de l'itération suivante. Un arbre de segmentation est donc le produit final de l'algorithme multirésolutions. L'exploration de toutes les échelles possibles assure la détection d'un motif répété, et ce, peu importe la taille de ce motif par rapport à la zone couverte par l'image. Suivant la segmentation un lot de descripteurs est extrait pour chacune des zones segmentées.

Ces descripteurs sont :

- le contraste moyen sur la frontière de la région étudiée;
- l'aire de la région étudiée, excluant l'aire des régions internes;
- l'écart type de l'aire des régions internes à la région étudiée;
- le vecteur de déplacement entre le centroïde de la région étudiée et le centroïde de la région parente;
- le périmètre de la région étudiée;
- le rapport entre les deux axes principaux de la région (calculés en trouvant la plus grande ellipse pouvant entrer dans la région);
- les coordonnées du centroïde de la région.

Note : Ces descripteurs sont soumis à une analyse par composantes principales (ACP) conservant 95 % de la variance.

Par la suite, l'estimation des modes de la fonction de densité de probabilité (PDF) des descripteurs est effectuée à l'aide d'une version modifiée de l'algorithme « meanshift ». Cet algorithme démarre en sélectionnant un point aléatoire dans l'espace des descripteurs. Par la suite, il visite une séquence de points qui permet à chaque itération de se déplacer dans la direction du gradient positif le plus fort dans l'espace de la densité. Cette procédure est exécutée en parallèle et est initialisée avec un ensemble de points de départ aléatoires.

Au terme de l'exécution, l'algorithme converge vers les maximums locaux dans l'espace de la densité. Ces maximums correspondent à un mode de la fonction de densité de probabilité, soit un regroupement de régions possédant des descripteurs semblables. Les régions qui peuvent être associées à un des modes trouvés par l'algorithme du « *meanshift* » sont considérées comme les texels qui appartiennent à une texture. Chaque mode correspond donc à une texture distincte de l'image, composée de multiples *texels* semblables.

Puisque l'espace des descripteurs possède de multiples dimensions, les données tendent à être très dispersées, l'algorithme a donc été modifié pour admettre un noyau à variance

variable. Ainsi, la magnitude du vecteur de déplacement devient plus grande lorsque les données sont dispersées et devient petite lorsque l'algorithme converge vers un mode de la fonction de densité, ce qui accélère le traitement et assure la convergence.

La procédure de recherche des modes de l'espace de densité par l'algorithme « meanshift » est généralement faite avec un noyau gaussien. Pour améliorer les résultats de segmentation, l'auteur propose une deuxième modification à l'algorithme : l'utilisation d'un noyau hiérarchique. Les *texels* ne sont pas toujours des zones uniformes et peuvent parfois être constitués d'une hiérarchie complexe de sous-régions. L'article suggère donc l'utilisation d'un noyau hiérarchique au sein de l'algorithme de segmentation. Cette modification est responsable d'une grande amélioration des résultats de segmentation comme en témoigne le paragraphe suivant.

L'algorithme proposé a été testé dans quatre environnements distincts. L'évaluation quantitative la plus détaillée présentée est basée sur un lot d'images générées artificiellement à l'aide des textures naturelles de Brodatz. Chacune des 100 images qui ont été segmentées est un collage de plusieurs textures naturelles disposées aléatoirement dans l'image. Les frontières entre les diverses textures sont représentées par des formes géométriques simples. L'erreur de segmentation est calculée simplement par le rapport entre le nombre de pixels erronés sur le nombre de pixels total des zones texturées. Globalement, l'algorithme arrive à segmenter correctement 93,3 % des pixels d'une image, en moyenne. Le tout soumis à une variance de 3,7 %. L'utilisation d'un noyau gaussien simple réduit les performances à 77,9 %  $\pm$  4,1. Ce qui justifie l'implantation du noyau hiérarchique. Enfin, 62,3 %  $\pm$  7,8 des pixels sont classifiés correctement lorsque l'algorithme original du « meanshift » est utilisé, soit sans le noyau hiérarchique et à variance variable.

Plusieurs autres tests ont été effectués avec la base d'images naturelles de Prague, de « Aerial-Produce » et de Berkeley. Les résultats de ces tests sont majoritairement qualitatifs... Les résultats présentés dans l'article démontrent que l'algorithme proposé est capable de surclasser d'autres propositions de la littérature. Toutefois, la complexité de

l'algorithme proposé exige un temps de calcul assez important. Il est évalué à 5 minutes pour une image de 512x512 pixels dans l'environnement MATLAB sur un ordinateur cadencé à 3.1 GHz.

Le même auteur a poursuivi la recherche sur les textures en composant un deuxième article d'intérêt sur la problématique (Ahuja et Todorovic, 2007). Cet article propose une méthode de segmentation des *texels* contenus dans une image numérique présentant une texture naturelle. Selon le texte, une texture est formée par une répétition spatiale d'un grand nombre de *texels*. Les *texels* sont des régions d'une image, dont les propriétés géométriques, photométriques et topologiques peuvent être caractérisées de manière stochastique. La méthode proposée permet d'isoler cet élément, soit le plus simple élément, pouvant être extrait d'une texture et permet même de le reconstruire, dans le cas où l'interocclusion des *texels* en cache partiellement les diverses parties.

Aucune des méthodes proposées par la littérature ne segmente précisément le *texel* d'une texture. Les approches actuelles utilisent des voisinages carrés, ronds ou d'une forme qui doit être définie au préalable par l'utilisateur afin de segmenter correctement les éléments composant une texture. Toutefois, la segmentation des *texels* doit être considérée et être utilisée comme base dans la segmentation d'images texturées. C'est pourquoi l'auteur propose une nouvelle méthode pour segmenter les *texels*. L'article a été écrit dans l'optique d'atteindre les buts suivants :

- identifier l'élément de base d'une texture; soit le *texel* d'une image texturée;
- obtenir un modèle générique complet de cet élément de base à partir des différents *texels*, incomplets (par cause d'occlusion) ou déformés;
- segmenter les *texels* en utilisant le modèle générique.

Pour isoler les *texels*, la texture est d'abord représentée par un arbre de segmentation dont la structure emmagasine la relation entre les diverses régions obtenues par une segmentation multirésolutions. Dans cette représentation, les *texels* apparaissent sous la forme de sous-

arborescences possédant une structure semblable. Les feuilles de l'arbre de cette structure possèdent des caractéristiques photométriques et géométriques similaires. L'extraction d'un modèle du texel complet, sans occlusion, est ensuite achevée à l'aide de l'opération d'union d'arborescence. Ce modèle complet est utilisé pour segmenter les *texels* de l'image qui peuvent être légèrement différents du modèle. Voici en détail ces diverses étapes :

Premièrement, la segmentation multirésolutions de l'image est effectuée selon la procédure présentée précédemment par Todorovic et Ahuja (2009).

Deuxièmement, l'algorithme tente de discerner les *texels*. L'identification des *texels* s'effectue en analysant l'arbre de segmentation généré précédemment. Pour localiser les arborescences qui correspondent à des *texels*, il suffit de trouver les sous arbres de la structure qui possèdent des caractéristiques topologiques, géométriques et photométriques similaires. Pour s'assurer que le lot d'arborescence choisi représente bel et bien des régions de l'ordre du texel et non du sous-texel, il suffit de trouver le plus grand groupement d'arborescences similaires de grande taille. Il faut donc chercher à optimiser deux critères, soit le nombre d'arborescences similaires dans le groupe et la taille (le nombre de niveaux) des arborescences utilisées pour former ce groupe. Selon l'auteur, il peut être pris pour acquis que le plus grand groupe d'arborescences de grande taille est un groupe de *texels* appartenant à une même texture. Pour évaluer la similarité des sous-arborescences, un algorithme d'appariement d'arbre est utilisé. L'algorithme d'appariement proposé par l'auteur est un autre travail qui a été exécuté précédemment par l'équipe. Cet algorithme calcule une mesure de distance basée sur la topologie et sur un certain nombre de critères choisis par l'utilisateur. Le lot de critères utilisé pour l'appariement d'arbre dans le contexte de segmentation des *texels* est évalué en comparant les caractéristiques de la région  $i$  par rapport à sa région parente  $j$ . Ce calcul relatif assure une invariance à la rotation et à l'échelle. Voici les caractéristiques utilisées :

- le contraste relatif, en niveaux de gris du texel  $i$  par rapport au parent  $j$ ;
- le ratio des aires entre la région  $i$  et sa région parente  $j$ ;
- le déplacement en  $x$  et en  $y$  du centroïde de  $i$  par rapport au parent  $j$ ;

- une mesure représentant la dispersion des aires des régions enfants de  $i$ ;
- un histogramme représentant la forme de la région  $i$ ;
- cet histogramme est calculé en comptant les pixels tombant entre 40 arcs tracés à partir du centroïde de la région vers l'extérieur. Par exemple, l'arc possédant le plus grand nombre de pixels contient l'axe principal de la région;
- l'angle entre l'axe principal de la région  $i$  par rapport à l'axe de la région  $j$ .

Troisièmement, on obtient un modèle générique du *texel* en fusionnant diverses arborescences. Le lot de *texels* détectés à l'étape précédente peut correspondre à un *texel* entier ou à un *texel* partiellement représenté. Pour fusionner les diverses arborescences de *texel*, l'auteur propose de trouver l'union des sous-arborescences en utilisant un réseau bayésien. Finalement, une segmentation des *texels* de l'image à l'aide du modèle générique obtenu à l'étape précédente est accomplie par l'algorithme d'*espérance maximisation*.

L'algorithme a été testé sur une base de données composée de 80 textures naturelles homogènes. Chaque classe de texture est composée de trois images de 320 par 240 pixels. L'auteur insiste sur la pertinence des données choisies puisqu'elles représentent un défi de taille pour l'algorithme de segmentation des *texels*. Par exemple, on indique que la base de données comprend de nombreux cas d'interocclusion des *texels*, de textures imbriquées dans les *texels*, de répétitions internes aux *texels* pouvant être confondus avec les *texels* eux-mêmes, de grandes variations entre l'apparence des *texels*, etc.

L'évaluation de la qualité des résultats a été effectuée en comparaison avec une segmentation manuelle des *texels* dans les images de la base de données. Ainsi par les méthodes habituelles il est possible de calculer l'erreur de segmentation en calculant les faux positifs, les faux négatifs et l'erreur totale. L'erreur de segmentation avoisine les 17 % dans le meilleur des cas, ou le nombre de variables cachées  $K$  paramétrant l'étape du réseau bayésien, est de 6. Les résultats présentés sont aussi évalués qualitativement. Les images segmentées sont très convaincantes malgré le haut taux d'erreur stipulé dans les résultats. L'auteur décrit qu'une bonne segmentation est obtenue même lorsque les frontières entre les divers *texels* sont

floues, et même lorsque de multiples interocclusions des *texels* forment des régions complexes dans l'image. Toutefois, cet algorithme est lent, il nécessite entre 10 et 25 minutes par images considérant l'exécution de code C sur un ordinateur cadencé à 2,8 GHz avec 2 Go de mémoire vive.

### **1.3.4 L'extraction de caractéristiques**

Suite au processus de segmentation, la zone image où se trouve véritablement le sujet d'intérêt est isolée. Les pixels appartenant à cette zone peuvent maintenant être analysés d'un point de vue de leur morphologie, de leur texture et de leur couleur en vue d'établir une relation entre ces caractéristiques et l'identité du spécimen. Divers outils mathématiques sont utilisés pour simplifier l'information véhiculée par des milliers de pixels en quelques descripteurs (caractéristiques) qui décriront les spécimens avec précision. Ces méthodes, issues de la littérature, seront présentées par les prochaines sections.

#### **1.3.4.1 Le traitement de l'histogramme des tons**

L'histogramme des tons de gris d'une image (ou des teintes de couleur) renferme de multiples informations sur l'image. L'histogramme est obtenu simplement en sommant la quantité de pixels d'une image pour chacun des tons de gris admis par la profondeur de couleur choisie. Des caractéristiques relatives à la texture, à la luminosité ou même à la couleur peuvent en être dérivées. Comme mentionné en 1.3.3.1, l'histogramme peut être considéré comme une fonction de densité de probabilité, ce qui ouvre la porte à de multiples outils d'analyse statistique. La liste des descripteurs présentée ici est inspirée de Gonzalez et Woods (2008) et des travaux de Levasseur (2008). Dans cette description,  $n$  est le nombre de niveaux de gris de l'image (la profondeur de couleur) et  $P_i$  est la probabilité d'obtenir un ton de gris d'intensité  $x_i$ .

- La *moyenne d'intensité lumineuse* se calcule par l'équation (1.5) :

$$\bar{x} = \sum_{i=0}^{n-1} P_i x_i \quad (1.5)$$

- L'*écart type l'intensité lumineuse* se calcule comme illustré par l'équation (1.6) :

$$\sigma = \sqrt{\sum_{i=0}^{n-1} P_i (x_i - \bar{x})^2} \quad (1.6)$$

- La *valeur R du second moment* de l'histogramme des tons de gris se calcule quant à elle comme illustré à l'équation (1.7). Le second moment  $m_2$  est la variance de population évaluée sur l'histogramme des tons de gris.

$$R = 1 - \left( \frac{1}{1 + \frac{m_2}{(n-1)^2}} \right) \mid m_2 = \sum_{i=0}^{n-1} P_i (x_i - \bar{x})^2 \quad (1.7)$$

- Le *coefficient d'asymétrie* de la distribution des tons de gris se calcule comme illustré en (1.8) :

$$skewness = \frac{\sum_{i=0}^{n-1} P_i (x_i - \bar{x})^3}{\left( \sum_{i=0}^{n-1} P_i (x_i - \bar{x})^2 \right)^{3/2}} \quad (1.8)$$

- Le *coefficient d'aplatissement* de la distribution des tons de gris se calcule comme illustré en (1.9) :

$$kurtosis = \frac{\sum_{i=0}^{n-1} P_i (x_i - \bar{x})^4}{\left( \sum_{i=0}^{n-1} P_i (x_i - \bar{x})^2 \right)^2} - 3 \quad (1.9)$$



- Le coefficient d'*uniformité* de la distribution des tons de gris par l'équation (1.10) :

$$uniformity = \sum_i \left( \frac{P_i}{\sum P_i} \right)^2 \quad (1.10)$$

- L'entropie de la distribution des tons de gris est obtenue à partir de l'équation (1.11) :

$$entropy = - \sum_i P_i \log_2 P_i \quad (1.11)$$

#### 1.3.4.2 La matrice de cooccurrence des tons de gris

La construction d'une matrice de cooccurrence des tons de gris sur une image permet de caractériser la texture par plusieurs statistiques. Les caractéristiques utilisées sont dérivées des travaux d'Haralick, Shanmugam et Dinstein (1973) sur cette matrice. Parmi les métriques utilisées on compte treize fonctions calculées à partir d'une matrice de dépendance des tons de gris  $P$  pour l'image  $I$  à traiter. Chaque élément de la matrice de dépendance  $P$  est fonction de quatre paramètres, le ton de gris de départ  $i$ , le ton de gris d'arrivée  $j$ , l'angle de voisinage  $\theta$  et la distance de voisinage  $d$ . Voici une brève présentation des calculs à effectuer pour extraire les caractéristiques liées à la matrice  $P$ .

Premièrement, il faut effectuer le calcul de la matrice de dépendance des tons de gris pour une distance  $d$  et un angle  $\theta$ . La matrice  $P$  de dépendance des tons de gris est une matrice qui compte les occurrences des passages d'un ton de gris  $i$  vers un ton de gris  $j$  au cours d'une distance  $d$  le long de l'angle  $\theta$ . La matrice est généralement calculée pour des paramètres de distances et d'angles fixes. Ces paramètres sont donc à optimiser pour maximiser les performances de discrimination. Suite au dénombrement des transitions de tons de gris, on obtient une matrice de dépendance spatiale des tons  $P$  qui couvre toutes les combinaisons de  $i$  et  $j$  possibles. Cette matrice est donc carrée et possède autant de lignes et de colonnes qu'il y

a de niveaux de tons de gris dans l'image considérée. Dans l'exemple qui suis, la matrice sera simplement de taille 4 par 4 puisque l'image considérée possède une profondeur de couleur de 2 bits. La Figure 1.7, tirée de l'article d'Haralick, présente un exemple simpliste de matrice  $P$ .

Voici par exemple comment obtenir les deux premiers éléments de la première ligne de cette matrice  $P$  en fonction de l'image  $I$  de la Figure 1.7 pour une distance  $d$  de 1 et un angle  $\theta$  de  $0^\circ$ . Le premier élément de  $P$ , le nombre 4, obtenu puisqu'on observe quatre transitions entre un pixel de ton 0 à un pixel de ton 0 considérant une distance  $d$  de 1 et un angle  $\theta$  de  $0^\circ$  degrés. Notez qu'une observation à l'angle  $0^\circ$  implique implicitement les transitions à l'angle  $180^\circ$  lorsque les deux tons de gris comparés sont le même (sur la diagonale de la matrice  $P$ ). Le deuxième élément de la matrice  $P$ , le nombre 2, est obtenu puisqu'on observe deux transitions d'un pixel de ton 0 à un pixel de ton 1 selon la distance et l'angle choisi. Inversement, le deuxième élément de la première colonne est également 2, puisqu'on observe autant de transitions de tons de gris d'un pixel de ton 1 à un pixel de ton 0, c'est la relation inverse.

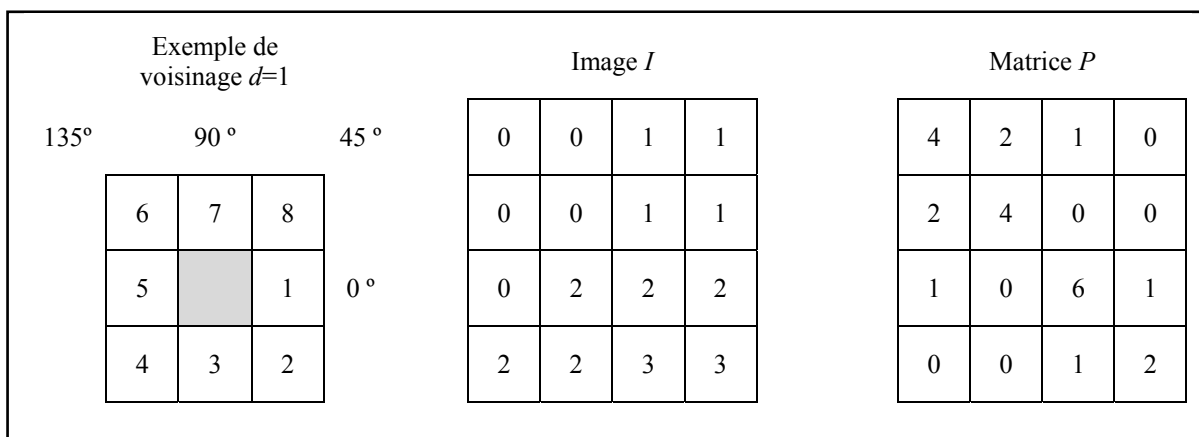


Figure 1.7 Exemple de matrice de cooccurrence avec  $d = 1$  et  $\theta = 0^\circ$ .

Adaptée d'Haralick, Shanmugam et Dinstein (1973)

Une fois la matrice de dépendance spatiale des tons de gris calculé, elle est normalisée de façon à ce que ses valeurs tiennent entre 0 et 1. Il est maintenant possible d'extraire diverses

statistiques qui permettent de caractériser de façon plus précise une texture donnée à partir de  $P$ . Les lignes suivantes présenteront une description très sommaire de ces statistiques. Premièrement une convention sera utilisée pour la présentation des fonctions. Voici la liste des variables utilisées :

- La variable  $p(i, j)$  réfère aux éléments de la matrice  $P$ ;
- La variable  $p_x(i)$  réfère à la probabilité marginale (à priori) obtenue en sommant les lignes de la matrice. Cette convention est généralisable pour  $p_y(j)$  qui somme les colonnes de la matrice. Enfin, d'autres variables  $p_{x+y}(k), p_{x-y}(k)$  réfèrent à la somme et à la différence des deux vecteurs  $p_x(i)$  et  $p_y(j)$ ;
- La variable  $N_g$  réfère au nombre de niveaux de gris utilisés dans l'image;
- Les variables  $\mu_x, \sigma_x$  réfèrent à la moyenne et à l'écart type respectivement mesurés pour le vecteur  $p_x(i)$ . Ces nomenclatures sont généralisables pour  $p_y(j)$  également;
- Les variables  $HX$  et  $HY$  réfèrent à l'entropie mesurée sur  $p_x(i)$  et  $p_y(j)$  respectivement;
- Les variables  $HXY1$  et  $HXY2$  sont calculées comme suis :

$$HXY1 = - \sum_i \sum_j p(i, j) \log\{p_x(i)p_y(j)\} \quad (1.12)$$

$$HXY2 = - \sum_i \sum_j p_x(i)p_y(j) \log\{p_x(i)p_y(j)\} \quad (1.13)$$

- La variable  $\epsilon$  réfère à un très petit nombre réel positif. Elle est utilisée pour éviter qu'un logarithme ne soit calculé pour la valeur 0, qui est non définie pour cette fonction.

Voici maintenant les fonctions qui permettent l'extraction des descripteurs :

- Le moment angulaire du second degré :

$$f_1 = \sum_i \sum_j (p(i, j))^2 \quad (1.14)$$

- Le contraste :

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left( \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \right) \quad (1.15)$$

- La corrélation :

$$f_3 = \frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (1.16)$$

- La variance :

$$f_4 = \sum_i \sum_j (i - \mu)^2 p(i, j) \quad (1.17)$$

- Le moment inverse :

$$f_5 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j) \quad (1.18)$$

- La moyenne de la somme :

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i) \quad (1.19)$$

- La variance de la somme :

$$f_7 = \sum_{i=2}^{2N_g} (i - f_6)^2 p_{x+y}(i) \quad (1.20)$$

- L'entropie de la somme :

$$f_8 = \sum_{i=2}^{2N_g} p_{x+y}(i) \log(p_{x+y}(i) + \epsilon) \quad (1.21)$$

- L'entropie :

$$f_9 = \sum_i \sum_j p(i, j) \log(p(i, j) + \epsilon) \quad (1.22)$$

- La variance de la différence :

$$f_{10} = \text{variance}(p_{x-y}) \quad (1.23)$$

- L'entropie de la différence :

$$f_{11} = - \sum_{i=2}^{2N_g} p_{x-y}(i) \log(p_{x-y}(i) + \epsilon) \quad (1.24)$$

- Les mesures de corrélation (2) :

$$f_{12} = \frac{f_9 - HXY1}{\max(HX, HY)} \quad (1.25)$$

$$f_{13} = [1 - \exp(-2.0(HXY2 - f_9))]^2 \quad (1.26)$$

### 1.3.4.3 Les motifs binaires locaux

Les descripteurs de textures basés sur les motifs binaires locaux ont initialement été proposés par Ojala, Pietikäinen et Harwood (1996). Le calcul de l'image résultante de l'application des motifs binaires locaux s'apparente à une opération de corrélation lors de l'application d'un filtre sur une image numérique. Il suffit de traiter chaque pixel de l'image en considérant les huit pixels de son voisinage immédiat (voisinage  $N_8$ ). Le voisinage d'un pixel forme une matrice de 3 par 3 pixels où le pixel à traiter est au centre et son voisinage est autour. La Figure 1.8 présente un exemple d'exécution de l'algorithme d'extraction des motifs binaires locaux relatif aux étapes présentées ci-dessous. Notez que cette présentation a été adaptée de Bouchard (2009a) .

1. Extraction du voisinage du pixel à traiter : Les huit valeurs de tons du voisinage du pixel à traiter sont extraites pour former une matrice de 3x3 pixels. Dans cet exemple, chaque pixel possède une valeur de ton de gris différente. Le pixel en cours de traitement possède la valeur de ton 4.
2. Un seuillage est effectué sur la valeur de ton des pixels du voisinage. Tout pixel possédant une valeur de ton supérieure ou égale à la valeur de ton du pixel en cours de traitement est attribué à la valeur 1. La valeur 0 est attribuée à toute valeur de ton inférieure à celle du pixel en cours de traitement.

3. Une matrice multiplicatrice est mise en mémoire. Cette matrice sera utilisée dans la prochaine étape de l'algorithme pour décrire la forme binaire locale résultante de façon unique.
4. Une multiplication élément par élément est effectuée entre la matrice résultante du seuillage de l'étape 2 et la matrice multiplicatrice de l'étape 3.
5. La sommation des valeurs de la matrice résultante de l'étape 4 est effectuée. Cette somme est rapportée dans l'image de sortie aux coordonnées correspondantes du pixel à traiter dans l'image d'entrée.
6. L'algorithme ré-exécute les étapes 1 à 5 jusqu'à ce que tous les pixels de l'image d'entrée soient traités.
7. Suivant la procédure d'identification des motifs binaires locaux, un histogramme est calculé pour caractériser la fréquence d'apparition des divers motifs. Le nombre calculé pour chaque pixel à l'étape 5 identifie de façon unique un motif de ton de gris parmi les motifs possibles. La forme de l'histogramme résultant est caractéristique de la texture étudiée par l'algorithme des motifs binaires locaux.

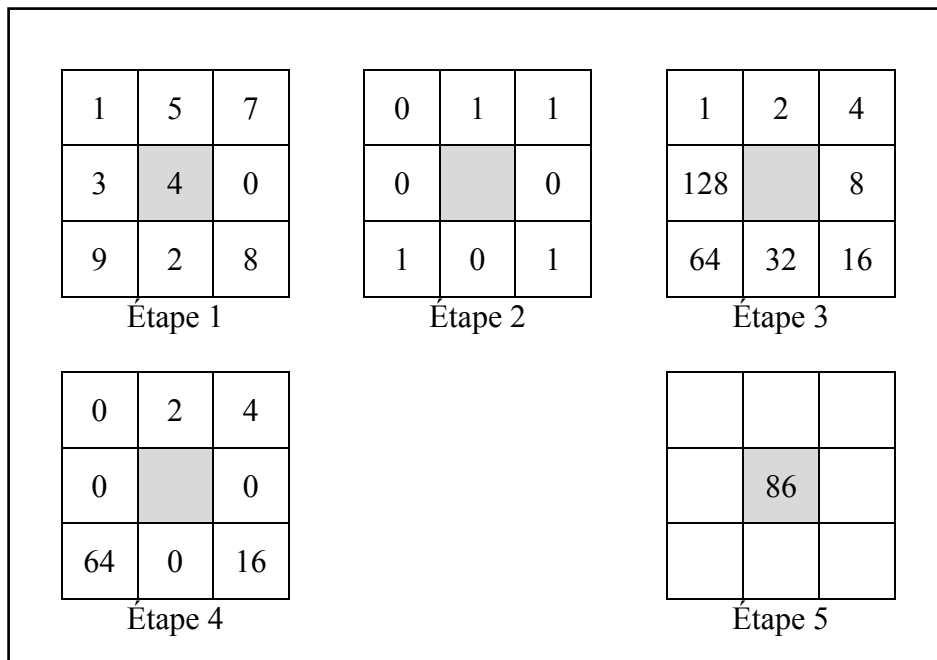


Figure 1.8 Exemple de filtrage par motifs binaires locaux.

L'algorithme permet aussi de calculer une distribution caractéristique du contraste de chaque forme binaire locale. Cette mesure produit un second histogramme où chaque valeur de pixel consignée est caractéristique de la mesure de contraste calculé comme suis : la moyenne des tons de gris des pixels du voisinage qui ont été assignés à la valeur 1 à l'étape 2 moins la moyenne des tons de gris des pixels assignés à 0. Pour l'exemple précédent, cette valeur de contraste est égale à 4,83 soit :  $(5 + 7 + 9 + 8)/4 - (1 + 3 + 0 + 2)/4$ .

Enfin, pour caractériser la texture par de simples descripteurs numériques à partir de la matrice des motifs binaires locaux, il suffit de calculer certaines statistiques sur l'histogramme des formes. Les moments de l'histogramme, dont les coefficients d'asymétrie et d'aplatissement sont couramment utilisés à cette fin.

#### **1.3.4.4 La transformée de Fourier**

La transformée de Fourier est utile pour faire une évaluation de la texture, par les composantes fréquentielles de l'image. Son calcul est une opération matricielle qui résulte en une représentation spectrale. Comme illustré à la Figure 1.9, dans cette représentation il est évident de repérer le contenu fréquentiel d'une texture. Dans la réponse à la transformée de Fourier centrée, l'information fréquentielle subit une rotation antihoraire de 90° par rapport à l'image originale. Au centre de la transformée centrée, on retrouve les basses fréquences et plus on s'éloigne du centre de la réponse plus l'information représentée est relative aux hautes fréquences de l'image.

Ces deux effets sont visibles dans l'exemple de la Figure 1.9 : au centre de la transformée de l'image originale on retrouve une grande quantité d'énergie fréquentielle, ce sont les basses fréquences : soit l'énergie attribuable à la luminosité moyenne, la composante continue. Du côté de la transformée de l'image altérée, on retrouve deux autres fréquences très énergétiques distribuées sur l'axe central vertical. Ces énergies fréquentielles représentées verticalement sur la transformée sont relatives à la trame fréquentielle horizontale qui a été superposée à l'image, ce qui démontre cette rotation de 90° de l'information.



Figure 1.9 Magnitude de la transformée de Fourier.  
Image originale adaptée de Gonzalez et Woods (2008)

Pour calculer cette transformation, un algorithme nommé la *transformée de Fourier rapide* est utilisé. Il existe de multiples implémentations de cette transformée qui ont des temps d'exécution différents en fonction de l'application désirée. L'implémentation que nous avons utilisée est celle mise à profit par MATLAB® tirée des travaux de Cooley et Tukey (1965). Cet algorithme utilise plusieurs optimisations pour calculer la transformée de Fourier discrète telle qu'illustrée par l'équation (1.27) qui transforme une matrice  $x_n$  avec un vecteur à  $d$  dimensions d'indice  $x_n = (n_1, n_2, \dots, n_d)$  par un lot de  $d$  sommations (sur  $n_j = 0 \dots N_j - 1$ ) où la division  $n/N$  correspond à  $\frac{n_1}{N_1}, \dots, \frac{n_d}{N_d}$ .

$$X_k = \sum_{n=0}^{N-1} e^{-i2\pi k(n/N)} x_n \quad \forall k = (0, \dots, N-1). \quad (1.27)$$



Plusieurs descripteurs peuvent être extraits de la réponse à la transformée de Fourier (Journaux *et al.*, 2008). Toutefois, la présentation des divers choix de descripteurs sera effectuée au cours des sections pertinentes de la méthodologie expérimentale.

#### **1.3.4.5 La réduction de dimensionnalité**

Maintenant que de multiples descripteurs ont été présentés, il ne reste qu'une étape (optionnelle) avant de passer à l'entraînement d'un algorithme de RDF. Chaque descripteur sélectionné sur un problème de classification augmente la capacité de discrimination interclasses. Toutefois, cette capacité de discrimination est contre balancée par la *malédiction de la dimensionnalité* décrite par Richard E. Bellman (Bellman, 1961). En effet, plus les données sont représentées dans un espace à grande dimension, plus il devient difficile de modéliser les distributions des classes, considérant un nombre fini d'échantillons. Ainsi, lorsque le nombre de descripteurs s'accroît, malgré l'augmentation de la capacité de discrimination, il peut arriver que les performances d'un classificateur chutent.

Pour limiter cette problématique, la réduction de dimensionnalité permet de sélectionner un nombre réduit de descripteurs tout en maximisant la capacité de discrimination de l'ensemble de description. Il suffit donc de valider quel ensemble de descripteurs est jugé pertinent dans ce compromis entre la discrimination et la facilité de modélisation. Une des méthodes populaires pour effectuer ce traitement est l'analyse par composantes principales (ACP).

L'ACP sélectionne les descripteurs pertinents en explorant les plus grands axes de variation des données (Duda, Hart et Stork, 2001). Après avoir décidé d'un nombre de descripteurs à retenir ou encore de la variance expliquée par le modèle qui doit être conservée, elle projette les données dans un nouvel espace où les principaux axes orthogonaux de variation sont exprimés par une combinaison linéaire des descripteurs originaux. Suite à l'exécution de la méthode, un ensemble réduit de descripteurs non corrélés est généré. Toutefois, ces descripteurs sont difficiles à interpréter, ils correspondent maintenant aux axes principaux de variation, sans nécessairement avoir une signification tangible en terme de description.

L'implémentation de cette méthode utilise les valeurs et vecteurs propres de la matrice de covariance sur les données. Les plus grandes valeurs propres correspondent aux axes de variation principaux dans les données. Suivant le tri en ordre décroissant de ces valeurs, les descripteurs originaux sont transposés dans l'espace ACP tel que décrit par l'équation (1.28) où  $eig = V$  est la matrice des vecteurs propres qui permet d'obtenir la matrice identité  $I$  à partir de la matrice  $C$ , selon  $V^{-1} * C * V = I$  et  $cov(x)$  est la matrice de covariance évaluée sur la matrice  $x$  qui représente les descripteurs dans l'espace euclidien original :

$$ACP_d = (x - \bar{x}) * eig_d(cov(x)) \quad (1.28)$$

### 1.3.5 L'optimisation de paramètres et l'entraînement

À cette étape, chaque spécimen échantillonné de la banque d'image est identifié par un point dans un espace euclidien à  $n$  dimensions. Les algorithmes de RDF permettent alors de modéliser la relation entre l'identité des spécimens et cet espace de description. Suite à l'entraînement de ces algorithmes de RDF, l'espace de description sera fractionné par un ensemble de frontières de décision ou modélisé par un ensemble de modes statistiques qui permettront de généraliser la relation entre les descripteurs et l'identité des spécimens. Cette section présentera les subtilités des divers algorithmes d'apprentissage statistiques qui seront mis à profit par la méthodologie expérimentale.

Notez que plusieurs de ces algorithmes de RDF admettent un lot de métaparamètres qui peuvent être optimisés indépendamment de la procédure d'apprentissage interne. Cette subtilité ne sera pas décrite ici, car il s'agit d'un choix d'implémentation spécifique. La méthodologie expérimentale présentée par les prochains chapitres inclura si nécessaire une section sur l'optimisation des métaparamètres. Notez que ce type d'optimisation est un vaste problème en soi. Certains algorithmes admettent un grand lot de métaparamètres ou une structure complexe de paramètres. Plusieurs méthodes avancées, tels les algorithmes évolutionnaires basés sur la théorie de Darwin (1837), peuvent être utilisées à cet effet.

### 1.3.5.1 Le classificateur naïf de Bayes

Le classificateur naïf bayésien (NB) modélise la distribution des données par un ensemble de gaussiennes dans un espace à  $n$  dimensions ( $n$  étant le nombre de descripteurs considérés). Ce classificateur prend donc pour acquis que les données sont distribuées de façon normale et ne possèdent qu'un seul mode par classe (ce qui lui vaut le qualificatif *naïf*). Pour entraîner un classificateur de Bayes, il suffit de calculer la moyenne et la covariance de la distribution des descripteurs pour chacune des classes  $c$  du problème considéré.

La généralisation s'effectue ensuite en évaluant pour chaque classe la fonction de densité de probabilité multivariée (PDF). Par la suite, la probabilité d'obtenir chaque classe  $c$  à partir des descripteurs d'un échantillon à classer est calculée à partir des différentes PDF. Enfin, la classe qui obtient la plus grande probabilité est attribuée à l'échantillon (Duda, Hart et Stork, 2001).

### 1.3.5.2 L'algorithme des $k$ plus proches voisins

L'algorithme des  $k$  plus proches voisins (kPPV) permet de classer un échantillon inconnu en comparant simplement la distance euclidienne (ou autre paradigme) entre cet échantillon et les spécimens connus dans l'espace de description. Cet algorithme non paramétrique ne s'exécute pas en deux phases d'apprentissage et de généralisation comme tous les autres. C'est également pour cette raison qu'il n'appartient à aucune des deux catégories d'algorithmes de RDF statistiques : les algorithmes de modélisation et ceux de séparation (modélisation des frontières). Certains le comparent à un algorithme de modélisation des centres de masse, puisque la distance calculée entre l'échantillon à classer et les données connues est inversement proportionnelle à la probabilité que l'échantillon à classer appartienne à l'une ou l'autre des classes des données connues. D'autres le comparent à un algorithme de modélisation de frontières, par sa dualité avec le diagramme de Voronoï (1908).

Considérant une base de données d'échantillons dont la classe est connue, la généralisation s'effectue en trouvant les  $k$  plus proches voisins de l'instance à classifier. Suivant l'identification des  $k$  voisins, un vote majoritaire sur la classe est tenu. Ainsi, si  $k = 3$ , advenant que deux voisins appartiennent à la classe 1 et qu'un seul voisin appartienne à la classe 2, l'échantillon sera classifié comme appartenant la classe 1. Le nombre de voisins est donc un paramètre qui peut être optimisé. Lorsqu'un nombre de voisins égal au nombre de classes est sélectionné, il peut arriver que l'algorithme ne converge vers aucune des classes. Un choix basé sur la distance doit être fait dans ce cas, ce qui revient à reformuler le problème avec une valeur de  $k = 1$ . Notez qu'il est aussi possible de considérer le type de distance comme un paramètre de l'algorithme. Toutefois, la distance euclidienne est généralement utilisée, nous ne présenterons donc pas les autres choix possibles.

### 1.3.5.3 L'arbre de classification C4.5

L'arbre de classification C4.5, nommé ainsi à cause du nom donné à l'algorithme pour le construire, est également un classificateur non paramétrique (qui n'utilise pas la modélisation statistique). Ce type de classificateur effectue une séparation parmi les données en prenant une série de décisions sur les descripteurs de ces données. Le but ultime est de construire un arbre de décision dont les feuilles sont « pures ». La pureté est obtenue lorsqu'en suivant le schéma décisionnel de l'arbre, chaque feuille converge vers une seule et unique classe parmi les données d'apprentissage.

Pour construire un tel arbre binaire, à chaque nœud, un élément important doit être considéré : le choix du descripteur utilisé pour prendre la décision. Pour ce faire, le C4.5 mesure le gain d'information associé à chaque décision possible et sélectionne le descripteur le plus approprié qui maximise ce gain  $\Delta i$  (Duda, Hart et Stork, 2001). Le gain d'information est basé sur la notion d'impureté  $i$  au nœud  $N$ , qui est évaluée à l'aide de l'entropie de Shannon, voir l'équation (1.29) où  $P(\omega_j)$  correspond à la fraction d'échantillons d'apprentissage qui appartiennent à la classe  $\omega_j$ .

$$i(N) = - \sum_j P(\omega_j) \log_2 P(\omega_j) \quad (1.29)$$

Le gain d'information mesuré par l'équation (1.30) soustrait de l'impureté au nœud courant  $i(N)$ , l'impureté restante au nœud descendant gauche  $i(N_L)$  et l'impureté restante au nœud descendant droit  $i(N_R)$ . Les impuretés à gauche et à droite sont pondérées par la fraction  $P$  des données d'apprentissage du nœud  $N$  qui iront au nœud gauche  $P_L i(N_L)$  et au nœud droit  $(1 - P_L) i(N_R)$ .

$$\Delta i(N) = i(N) - P_L i(N_L) - (1 - P_L) i(N_R) \quad (1.30)$$

Le procédé itératif de construction de l'arborescence s'arrête lorsque l'erreur de validation atteint son minimum. Une partie des données d'apprentissage est donc réservée pour valider l'arrêt de l'apprentissage. Dès que l'erreur de validation commence à augmenter, la construction de l'arborescence est stoppée, sans quoi, un sur apprentissage aurait lieu. Suite à l'apprentissage, une procédure d'émondage peut être exécutée pour simplifier l'arborescence et éliminer les décisions redondantes, toutefois nous ne couvrirons pas les détails sur cet algorithme ici. Plus d'informations sont disponibles dans l'ouvrage de Quinlan (1993). L'avantage d'un tel algorithme d'apprentissage réside dans son modèle en arborescence qui possède un sens tangible au sein de la problématique où il est appliqué. Les règles de décision concrètes peuvent être réutilisées pour analyser davantage le problème.

#### 1.3.5.4 Le réseau de neurones à base radiale

Contrairement au réseau de neurones de type perceptron, le réseau à base radiale (RBR) tente de modéliser les données par un ensemble de modes gaussiens, et non par modélisation des frontières de décision. Au départ, l'utilisateur doit configurer l'algorithme en sélectionnant le nombre de centres (modes) et la dispersion acceptée par ceux-ci. Le nombre de centres correspond au nombre de neurones sur la couche cachée du réseau, et la dispersion contrôle l'aspect de la fonction d'activation gaussienne de ces neurones cachés. La couche de sortie

de ce réseau neuronal est identique au perceptron : une combinaison linéaire de la sortie des neurones cachés.

Au cours de la phase d'apprentissage, cet algorithme optimise la position, la dispersion des noyaux et les poids utilisés pour faire la combinaison linéaire de la couche cachée. L'avantage de cette topologie est l'indépendance des couches lors de l'entraînement. La position et dispersion des noyaux peut être optimisée indépendamment des poids synaptiques ce qui simplifie beaucoup le problème (Witten et Frank, 2005a). Pour optimiser la position des noyaux à base radiale (RBF), un algorithme de partitionnement des données (*clustering*) peut être utilisé tel le *k-means* (MacQueen, 1967). Les poids synaptiques sont ensuite optimisés à l'aide d'une régression logistique qui minimise l'erreur de généralisation. L'équation (1.31) démontre cette optimisation du vecteur des poids synaptiques  $W$  appliqués à la réponse de la couche cachée  $\varphi$  qui résulte en la sortie  $z$  pour un vecteur d'entrée  $x$ . Il suffit de calculer la matrice pseudo-inverse  $\varphi^\dagger$  et de la multiplier au vecteur de sortie désiré  $T$  pour obtenir le vecteur poids  $W$  nécessaire à l'optimisation (Duda, Hart et Stork, 2001).

$$z(x) = W\varphi \rightarrow W^t = \varphi^\dagger T \quad (1.31)$$

### 1.3.5.5 Le séparateur à vaste marge

Le séparateur à vaste marge (SVM), aussi nommé machine à vecteurs de support (Cortes et Vapnik, 1995), est un algorithme de RDF qui s'intéresse à la modélisation des frontières de décision. Plus précisément, le critère d'optimisation est basé sur la maximisation de la marge qui sépare la frontière de décision des classes, ce qui admet une meilleure capacité de généralisation. Une transformation est utilisée au cours de la phase d'entraînement : les données sont projetées dans un espace de plus grande dimension, dans lequel elles pourront plus facilement être séparées linéairement. Toutefois, la description des données dans cet espace augmenté complexifie énormément le problème. Pour éviter cette complexification, l'« astuce du noyau » est utilisée (*kernel trick*). Ainsi, l'hyperplan optimal de séparation peut être calculé sans avoir à projeter les données et analyser la séparation dans l'espace

augmenté. Puisqu'un seul hyperplan est considéré par le SVM, il s'agit d'un classificateur binaire. Il est donc nécessaire d'utiliser plusieurs SVM pour traiter un problème qui considère plus de deux classes. La section 1.3.6 traitera des astuces possibles pour faire la fusion des résultats de multiples SVM.

L'hyperplan de décision optimal est calculé de façon à maximiser la marge entre les données de deux classes à séparer. La position de l'hyperplan optimal est fonction des données d'apprentissage de chaque classe qui sont les plus proches les unes des autres. Ces données sont situées de part et d'autre de l'hyperplan optimal et sont nommées « vecteurs de support ». La fonction de décision liée à cet hyperplan s'exprime simplement par  $f(x) = wx + b$  où  $f(x) \geq 0$  lorsque l'échantillon  $x$  appartient à la classe  $y = 1$  et  $f(x) < 0$  lorsque l'échantillon  $x$  appartient à la classe  $y = -1$ . Considérant ce système d'inéquations et un problème séparable linéairement, il existe au moins un hyperplan qui vérifie l'inéquation  $f(x_i)y_i > 0 \forall i \in \{1, \dots, n\}$ , où  $x_i$  est le vecteur de descripteurs de l'échantillon  $i$  et  $y_i$  est la classe de cet échantillon (avec  $y_i \in \{-1, 1\}$  puisqu'on traite d'un problème binaire). Dans ces conditions, il suffit de minimiser l'équation (1.32) pour découvrir l'hyperplan optimal (Milgram, 2007). Notez que la portion de droite de l'équation admet deux nouveaux termes liés au traitement de données non séparable linéairement : la variable d'écart  $\xi_i$  et le facteur de régularisation  $C$  qui détermine la pénalité d'une donnée mal classée dû à l'impossibilité d'une séparation linéaire.

$$\min_{i=1, \dots, n} \{|wx_i + b|\} = 1 \Leftrightarrow \min_{i=1, \dots, n} \left\{ \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \right\} \mid y_i(wx_i + b) \geq 1 - \xi_i \quad (1.32)$$

Ce système d'équations est généralement résolu en utilisant les multiplicateurs de Lagrange. Pour plus de détails sur la mathématique nécessaire, la thèse de Milgram (2007) et l'ouvrage de Duda, Hart et Stork (2001) présentent un bon résumé de la situation, ainsi que des subtilités de l'astuce du noyau. Notez que dans sa description complète, le SVM admet plusieurs métaparamètres qui doivent être fixés par son utilisateur. Le facteur de régularisation  $C$  doit être optimisé pour chaque problème. De plus, le choix du type de noyau

(typiquement linéaire, polynomial, à base radiale ou encore de forme sigmoïde) est un autre métaparamètre important qui doit être fixé (et non optimisé). Le choix de noyau implique aussi plusieurs autres métaparamètres puisqu'il impose un lot de constantes (biais et degré), qui doivent également être optimisées ou fixées en fonction du problème à résoudre.

### 1.3.5.6 Le perceptron multicouche

Le perceptron multicouche (PM) est un réseau de neurones qui tente de positionner un ensemble de frontières de décision dans l'espace de description pour modéliser la problématique. Cet algorithme est sans aucun doute le plus complexe à utiliser, considérant ceux présentés par notre étude. Le tout, puisqu'il admet une grande flexibilité et donc un lot important de paramètres à optimiser. La topologie est sans aucun doute le critère le plus déterminant de la performance. Avec le perceptron, il est possible de configurer la topologie par le nombre de couches cachées  $h_j$  ainsi que par le nombre de neurones sur chacune des couches. Plus le nombre de neurones cachés est grand, plus la capacité de modélisation s'accroît. Il devient alors possible de modéliser des relations mathématiques très complexes.

L'entraînement d'un réseau perceptron est effectué par rétropropagation de l'erreur de classification (voir Figure 1.10). Considérant un vecteur de description  $X_k$  présenté à l'entrée du réseau, et la classe désirée  $d_m$  de l'échantillon (sous la forme d'un vecteur de longueur  $m$ ) présenté à la sortie  $y_m$  du réseau; les poids synaptiques des neurones (vecteurs  $v$  et  $w$ ) sont mis à jour de façon à minimiser l'erreur. Ce procédé est piloté par la descente de gradient dans l'espace des poids synaptiques. Le processus itère en présentant successivement les  $k$  données d'apprentissage au réseau. Chaque fois que la base de données d'apprentissage est présentée complètement, une « époque » est achevée. À chaque époque, le taux d'erreur observé tend à diminuer. L'entraînement est arrêté lorsque l'erreur de validation (évaluée sur une base de données indépendante de l'entraînement) cesse sa décroissance. Lorsque ce phénomène se produit, le réseau de neurones est bien entraîné. Si l'entraînement est poursuivi passé ce nombre d'époques, les données d'entraînement seront mémorisées



intégralement par le réseau, qui perdra alors sa capacité de généralisation. Ce phénomène est appelé le surapprentissage.

**0.** L'initialisation des poids synaptiques  $w$  (entrée) et  $v$  (sortie) est obtenue en tirant au hasard selon une loi uniforme sur un intervalle  $[-\varepsilon, +\varepsilon]$  avec  $\varepsilon \geq 1$ .

**1.** Pour  $k \in \{1, 2, \dots, K\}$  (parcours de la base d'apprentissage)

1.1 La forme  $X_k$  est présentée à l'entrée du réseau

1.2 Une propagation directe dans le réseau de cette forme est réalisée \*:

$$h_j(k) = f\left(\sum_{n=0}^N v_{jn} x_n(k)\right) \quad y_m(k) = f\left(\sum_{j=0}^J w_{mj} h_j(k)\right)$$

1.3 L'erreur d'époque est calculée :

$$E = E + E(k) = E + \frac{1}{2} \sum_{m=1}^M (d_m(k) - y_m(k))^2$$

1.4 Le signal d'erreur de sortie  $\delta_m(k)$  est calculé \* :

$$\delta_m(k) = (d_m(k) - y_m(k)) f'\left(\sum_{j=0}^J w_{mj} h_j(k)\right)$$

$$\delta_m(k) = (d_m(k) - y_m(k)) y_m(k) (1 - y_m(k))$$

1.5 La rétropropagation de  $\delta_m(k)$  pour obtenir le signal d'erreur  $\delta_j(k)$  associé à chacun des neurones  $j$  de la couche cachée \* :

$$\delta_j(k) = h_j'(k) \sum_{m=1}^M w_{mj} \delta_m(k) = h_j(k) (1 - h_j(k)) \sum_{m=1}^M \delta_m(k) w_{mj}$$

1.6 Le poids des connexions parvenant aux couches cachée et de sortie est modifié :

$$\Delta w_{mj} = \eta \delta_m(k) h_j(k) \quad \Delta v_{j,n} = \eta \delta_j(k) x_n(k)$$

**2.** Mise à jour du compteur d'époques :  $t = t + 1$

Calcul du taux d'erreur :  $E = \frac{E}{K}$

Appliquer la condition d'arrêt si désiré, sinon aller à 1.

\* Notez que la fonction d'activation  $f$  est une sigmoïde dont la dérivée est  $y' = f'(a) = y(1 - y)$ , d'où les équivalences proposées en 1.4 et 1.5.

Figure 1.10 Algorithme de la rétropropagation de l'erreur.  
Adaptée de Lepage et Solaiman (2003)

### **1.3.6 La généralisation**

Enfin, la dernière étape nécessaire à la mise en place d'un système intelligent de RDF dans les images numériques est la généralisation. C'est à cette étape qu'on utilise le modèle mathématique de reconnaissance de forme, précédemment entraîné, pour classer de nouvelles données jamais vues par le système. Toutefois, il ne s'agit pas encore de la mise en service du système comme telle. Avant de procéder, plusieurs autres calculs peuvent être effectués pour traiter la sortie des algorithmes de RDF tels la fusion des décisions, le rejet d'une donnée et l'évaluation quantitative des performances.

#### **1.3.6.1 La fusion des décisions**

La fusion de la décision est nécessaire lorsque plusieurs classificateurs entrent en jeu. Il existe plusieurs façons de fusionner l'information dans un système. Nous nous intéresserons ici qu'à la fusion de l'information qui se situe après la classification dans la chaîne de traitement. Notez donc qu'il est possible de faire la fusion de diverses modalités, par exemple en combinant l'information de l'image et du son, ou encore au niveau des descripteurs en combinant par exemple l'information de la texture et de la couleur et enfin au niveau des décisions, en combinant la réponse de plusieurs outils de classification (Bouchard, 2009d; Ross et Jain, 2003).

Dans la section traitant du SVM, nous avons abordé la problématique de ce classificateur binaire dont la sortie n'admet que deux classes possibles. Pour pouvoir traiter un problème avec plus de deux classes, l'utilisation de plusieurs SVM est nécessaire. Dans ce contexte, deux paradigmes de fusion simples sont souvent utilisés : la fusion un-contre-tous, ou la fusion un-contre-un. En mode un-contre-tous, il y aura un SVM entraîné pour chaque classe du problème. Ainsi, pour un problème à 5 classes, il y aura 5 SVM, dont la tâche est de départager les échantillons d'une des classes de la distribution des échantillons de toutes les autres classes. Le classificateur possédant la plus forte réponse en sortie (le score) sera sélectionné. L'échantillon sera alors attribué à la classe traitée par ce classificateur.

En mode un-contre-un, chaque SVM est entraîné à différencier deux des classes en cause. Ainsi, pour un problème à 5 classes, il y aura autant de classificateurs à entraîner que groupes de deux classes possibles, soit  $5 \times 4 \div 2 = 10$ . Suite à la présentation d'un échantillon au système de classification, la classe étant la plus favorisée par le système est assignée à l'échantillon suite à un vote majoritaire. Cette méthode est recommandée et implémentée par les concepteurs de la librairie LibSVM (Chang et Lin, 2001), puisqu'elle implique moins de temps de calcul par son nombre réduit de vecteurs de support (Chih-Wei et Chih-Jen, 2002).

### 1.3.6.2 Le mécanisme de rejet

Le rejet est une option intéressante lorsque l'erreur de classification implique des conséquences graves. Ce mécanisme permet de détecter les échantillons passés au système de classification dont la confiance sur la décision n'est pas assez élevée. Ces échantillons rejetés peuvent alors être envoyés au traitement manuel. Ce mécanisme est simple à implémenter, il suffit d'appliquer un seuil à la réponse du classificateur (le score de classification). Chaque paradigme de classification est capable de fournir un score, une mesure qui indique avec quelle précision l'échantillon classifié respecte le modèle de la classe tel qu'établi par le classificateur. Deux méthodes de rejet différentes ont été étudiées : celle proposée par Chow (1970), qui utilise un seuil global de décision pour toutes les classes et celle proposée par Fumera (2000), qui utilise un seuil adapté à chaque classe.

1. Pour chaque seuil  $t$  entre 0 et 1 avec incréments de 0,001
  - a. Calculer le taux d'erreur  $e(t)$  sur la base d'apprentissage
  - b. Calculer l'erreur normalisée  $e_n(t) = e(t)/\max(e(t))$
2. Trouver le seuil qui minimise la fonction de coût  $f_c = e_n(t) + r(t)$ .

Figure 1.11 Algorithme d'ajustement des seuils de Chow.  
Adapté de Chow (1970)

La méthode de Chow (voir Figure 1.11) est simple à mettre en place, il suffit d'utiliser un lot de données pour valider le seuil le plus efficace lorsqu'on met en relation le taux d'erreur et

le taux de rejet (selon le critère de notre choix, en fixant le taux d'erreur maximal toléré, ou encore le taux de rejet maximal toléré pour un traitement manuel).

Toutefois, étant donné une base de données d'entraînement finie, un léger déphasage entre la distribution échantillonnée et la population réelle peut favoriser le rejet d'une classe au détriment d'une autre. Ce phénomène entraîne une augmentation du taux de rejet, pour une performance fixée à un taux d'erreur donné. Pour diminuer l'impact du biais d'échantillonnage, la méthode de Fumera (voir Figure 1.12) propose l'utilisation d'un ensemble de seuils (un seuil de rejet pour chaque classe). Avec l'ensemble de seuils optimal, Fumera démontre mathématiquement que le taux de rejet pour un taux d'erreur fixe sera toujours inférieur ou égal en comparaison avec la méthode proposée par Chow. Cependant, l'optimisation d'un ensemble de seuils est un problème très complexe en soi. Un algorithme holistique (supérieur à celui présenté par la Figure 1.12) doit être utilisé pour résoudre ce problème, tel un algorithme évolutionnaire.

1. Initialiser le vecteur  $\vec{t}_c$  avec les plus hauts seuils qui résultent en un rejet nul
2. Pour chaque taux de rejet  $r$  entre 0 et 1 avec incréments de 0,001
  - a. Initialiser  $c = 1$
  - b. Tant que le taux de rejet  $r$  n'est pas atteint
    - i. Augmenter le seuil  $t_c$  de la classe  $c$  de 0,001
    - ii. Calculer le taux de rejet global  $r(\vec{t}_c)$  sur la base d'apprentissage
    - iii. Changer de classe pour la prochaine itération de « b. » :  $c = c + 1$
    - iv. Si  $c \geq \text{nombre de classe}$  Alors  $c = 1$
  - c. Calculer l'erreur  $e(\vec{t}_c)$  pour ce point d'opération
3. Calculer l'erreur normalisée  $e_n(\vec{t}_c)$  pour tous les niveaux de rejets testés
4. Trouver l'ensemble de seuils qui minimise la fonction de coût  $f_c = e_n(\vec{t}_c) + r(\vec{t}_c)$ .

Figure 1.12 Algorithme d'ajustement des seuils de Fumera.  
Adapté de Fumera, Roli et Giacinto (2000)

### 1.3.6.3 Les matrices de confusion

Enfin, pour bien analyser les performances d'un système de classification plusieurs outils sont utilisés. La matrice de confusion est la base de plusieurs évaluations statistiques des performances, elle nous renseigne sur la distribution de l'erreur de classification interclasses. Pour construire cette matrice, il suffit de produire un tableau qui met en relation le nombre d'échantillons bien classés et mal classés. Le Tableau 1.5 présente un exemple de cette matrice.

Tableau 1.5 Exemple de matrice de confusion  
Adapté de Wikipedia (Wikimedia Foundation, 2010)

		Classe réelle		
		Chat	Chien	Lapin
Classe estimée	Chat	5	2	0
	Chien	3	3	2
	Lapin	0	1	11

Dans cet exemple, la base de données de test contient 8 chats, 6 chiens et 13 lapins. Toutefois, le classificateur a classé les données et indique qu'il y a 7 chats, 8 chiens et 12 lapins. La matrice de confusion permet de visualiser à quel endroit l'erreur de classification a lieu. On remarque que le classificateur reconnaît bien les lapins, toutefois il existe une grande confusion entre les chiens et les chats.

Diverses statistiques importantes peuvent être extraites de ces matrices. Les équations suivantes présentent ces informations (Witten et Frank, 2005a):

- Le taux de classification :

Le taux de classification calcule le rapport entre le *nombre d'échantillons correctement classifiés* (situés sur la diagonale de la matrice de confusion) et le *nombre total d'échantillons considérés*, comme illustré par (1.33). Cette mesure indique la performance globale du classificateur. Toutefois, elle n'est pas très pratique lorsque les classes considérées sont très déséquilibrées. Par exemple, un classificateur considérant 990 échantillons de la classe A et 10 échantillons de la classe B obtiendra 99% de taux de classification simplement en classifiant toutes les données dans la classe A. Toutefois, ce n'est pas un bon classificateur, puisqu'il est incapable de départager A de B. Le taux de classification de notre exemple au Tableau 1.5 est de 85,2 %.

$$\frac{\text{échantillons correctement classifiées}}{\text{nombre total d'échantillons}} \quad (1.33)$$

- La précision de classe :

La mesure de précision de classe permet de résoudre le problème du taux de classification global. Elle permet d'étudier le taux de confusion entre les classes après classification du problème. Cette mesure se calcule par le rapport entre le *nombre d'échantillons correctement classifiés de la classe désirée* (soit un des nombres de la diagonale de la matrice de confusion) et le *nombre total d'échantillons classifié vers cette classe* (soit la somme de la ligne correspondante de la matrice), comme illustré par (1.34). Dans notre exemple du Tableau 1.5 la précision sur les chats est de 71,4 %, sur les chiens est de 37,5 % et sur les lapins est de 91,7 %.

$$\frac{\text{échantillons correctement classifiées de classe C}}{\text{nombre total d'échantillons classifiés vers la classe C}} \quad (1.34)$$

- Le rappel de classe :

La mesure de rappel de classe permet d'étudier le rapport entre le *nombre d'échantillons correctement classifiés de la classe désirée* (soit un des nombres de la diagonale de la matrice de confusion) et le *nombre total d'échantillons classifié vers cette classe* (soit la

somme de la ligne correspondante de la matrice), comme illustré par (1.34). Dans notre exemple du Tableau 1.5 le rappel sur les chats est de 62,5 %, sur les chiens est de 50,0 % et sur les lapins est de 84,6 %.

$$\frac{\text{échantillons correctement classifiées de classe } C}{\text{nombre total d'échantillons véritablement de la classe } C} \quad (1.35)$$

#### 1.3.6.4 Les intervalles de confiance

Pour finalement évaluer la variation admise par les résultats de multiples répliques d'une expérience, les intervalles de confiance de Student (1908) peuvent être mis à l'emploi. Ce type d'intervalles de confiance est adéquat pour mesurer la moyenne d'un échantillon de faible nombre (typiquement où  $n < 30$ ). Pour calculer un intervalle de confiance sur la moyenne  $\bar{x}$  d'un certain échantillon de taille  $n$  il faut d'abord fixer le degré de confiance  $1 - \alpha$  qui détermine la précision de la phrase.

Considérant que la distribution des échantillons suis une loi de Student  $t$ , un taux de confiance typique de 95 % sur un intervalle de confiance implique que 95 % du temps, la moyenne de l'échantillon évalué sera fidèle à la moyenne de population lorsqu'elle se situe dans cet intervalle de confiance. Le calcul de l'intervalle de confiance implique aussi un autre paramètre, le degré de liberté  $\nu = n - 1$  qui est fonction du nombre d'échantillons. Considérant ces deux paramètres de distribution, le calcul de l'intervalle de confiance sur la moyenne  $\bar{x}$  d'un échantillon de variance  $s$  s'effectue comme illustré par l'équation (1.36).

$$ic = \bar{x} \pm t^{-1} \left( 1 - \frac{\alpha}{2}, n - 1 \right) \frac{s}{\sqrt{n}} \quad (1.36)$$

#### **1.4 Discussion sur la revue de littérature**

Ce chapitre a permis de présenter un portrait des travaux existants en reconnaissance automatisée du corail, des travaux antérieurs à ce projet et des méthodes de RDF utiles à cette problématique. L'étude de la littérature a révélé que peu d'auteurs se penchent actuellement sur la problématique automatisée de reconnaissance du corail. Sans que la littérature traite concrètement de ce problème, il semble que le faible nombre de travaux pertinents soit relié à l'absence de base de données publique pour encourager les échanges entre les chercheurs. En effet, les domaines très actifs de la RDF telle la reconnaissance de visages, de caractères manuscrits, de signatures, etc.; sont tous étudiés à l'aide d'un lot de bases de données publiques, ce qui n'est pas le cas pour la reconnaissance de spécimens coralliens. Les travaux similaires démontrent toutefois qu'il est possible de classifier automatiquement les spécimens de corail dans certaines conditions.

Les travaux effectués antérieurs à ce mémoire, sous la supervision du Dr Landry, ont quant à eux fixé le cadre logiciel à la base de l'étude. Même si les performances des systèmes proposés ne sont pas concluantes, le schéma de base d'un système de RDF y a été implémenté. Il s'agit donc d'une contribution notable. Enfin, l'étude de la littérature de la VA et de l'IA nous a renseignés sur les méthodes d'échantillonnage et de validation à favoriser, les outils de segmentations existants, les façons de décrire simplement la texture, la couleur et la morphologie d'un objet, les subtilités des algorithmes de RDF et l'interprétation des résultats produits par ces algorithmes. Les deux prochains chapitres présenteront de quelle façon tous ces éléments théoriques peuvent être mis à contribution dans un système complet de reconnaissance automatisé de spécimens coralliens à partir d'images numériques.



## CHAPITRE 2

### APPROCHE DE DESCRIPTION PAR FORMES BENTHIQUES

Les expérimentations basées sur les formes benthiques ont été rendues possibles grâce à la base de données fournie par nos collaborateurs de l'Université De La Salle, aux Philippines. Cette base de données contient les photographies et l'identification des formes de vies du benthos d'un site d'intérêt de la mer de Sulu-Sulawesi et qui ont été consignées au cours d'une expédition en mai 2006. Ces données sont pour nous d'un grand intérêt, par le vaste contenu historique qui pourrait éventuellement nous être accessible. En effet, ce site a été sujet à de nombreux échantillonnages photographiques entre 1996 et 2006 par l'équipe du Dr Licuanan dans le but d'évaluer l'évolution de la santé des organismes aquatiques qui y vivent. Puisque nous envisageons éventuellement d'automatiser complètement le suivi de l'intégrité écologique d'un point de vue de la diversité biologique, cette collaboration et l'expérimentation sur les données qui en découle sont d'une grande importance.

Les diverses sections de ce chapitre introduiront la nature des formes benthiques, le protocole d'acquisition des images et la base de données utilisée. Ensuite, les diverses expériences qui ont été tentées seront présentées. Pour assurer la clarté de l'exposé, la méthodologie, les résultats et la discussion seront présentés séquentiellement pour chacune de ces expériences. Enfin, une discussion générale sur la faisabilité d'une approche descriptive par formes benthiques conclura le chapitre.

*Notez que le point de départ de cette étude réside en divers travaux effectués par d'autres étudiants au préalable. La section 1.2 de la revue littéraire résume ces travaux et détaille le contexte méthodologique qui a inspiré les développements présentés par ce chapitre. De plus, quelques outils logiciels développés au cours de ces travaux ont été réutilisés ou modifiés pour réaliser les expériences présentées dans cette section.*

## 2.1 Introduction aux formes benthiques

Les formes benthiques offrent une identification simplifiée des coraux et des autres formes de vie aquatiques qui permet aux non-initiés d'effectuer facilement la saisie des données. Ces formes sont attribuées suivant l'apparence globale des colonies coralliennes et donnent aux biologistes un aperçu de l'état de santé de celles-ci. Introduites initialement par Susan Anne English *et al.* (1994), elles ont été reprises par l'*Australian Institute of Marine Science* (AIMS) dans leurs procédures standardisées de suivi des communautés benthiques. C'est la version légèrement modifiée de l'AIMS (Page *et al.*, 2001) qui a inspiré la structure la base de données d'images sous-marines des Philippines.

La description par formes benthiques est formée d'une hiérarchie à six niveaux qui décrivent plus ou moins précisément le spécimen observé. Au choix de la personne identifiant les spécimens, selon la qualité de l'image ou de sa connaissance au sujet de la taxinomie des coraux, l'identification peut prendre la forme d'un groupe benthique, d'une forme de vie benthique, d'une famille, d'un genre, d'une espèce ou d'un type de substrat marin (énumérés du moins précis au plus précis). À ce sujet, le document de l'AIMS cité plus tôt fournit une clé dichotomique (arbre décisionnel) qui dicte dans quelles conditions un objet peut être classé parmi un des groupes ou une des formes benthiques. La classification du point de vue taxinomique (famille, genre ou espèce) à partir de photographies requière une bonne expérience dans le domaine et est dans la grande majorité des cas non disponible dans notre base de données. À titre de référence, le Tableau 2.1 présente les formes benthiques ainsi que les types de substrats marins utilisés au sein de la base de données. Notez que ce tableau n'inclut pas les groupes benthiques puisque ce niveau grossier de classification n'a pas été utilisé dans la base de données.

Les formes benthiques du Tableau 2.1 évoquent plusieurs formes qui peuvent sembler familières. Elles le sont effectivement : les images de la Figure 2.1 montrent à quel point peuvent être évidentes les formes de ramifications (*branching*), digitées (*digitate*), encrustantes (*encrusting*), tabulées (*tabulate*), de goupillons (*bottlebrush*), foliacées (*foliose*),

massives (*massive*), sous-massives (*sub-massive*) et de champignons (*mushroom*). Cette simplicité de concept est au cœur de l'approche, car elle permet à tous de participer à la tâche colossale d'identification des images.

Tableau 2.1 La codification des formes benthiques

Code	Description	Code	Description
AA	Algal assemblage	CMR	Mushroom coral
ACB	Branching acropora	CS	Submassive non-acropora
ACD	Digitate acropora	DC	Dead coral
ACE	Encrusting acropora	DCA	Dead coral with algae
ACS	Submassive acropora	HA	Halimeda
ACT	Tabulate acropora	IND	Indeterminate
ACX	Bottlebrush acropora	MA	Macroalgae
CA	Coralline algae	OT	Other
CB	Branching non-acropora	R	Rubble
CE	Encrusting non-acropora	S	Sand
CF	Foliose non-acropora	SC	Soft coral
CHL	Heliopora (blue coral)	SI	Silt
CM	Massive non-acropora	SP	Sponge
CME	Millepora (fire coral)	TA	Turf algae

*Note : Les descriptions des formes benthiques ont été laissées en anglais pour éviter de trahir le sens de leurs abréviations.*

Au sujet des formes benthiques présentées à la Figure 2.1, on dénote une distinction marquée entre les coraux du genre *acropora* et les autres. Selon Licuanan (2001), on distingue un corail de genre *acropora* par son évidente structure coralliaire (*corallite*) axiale au sommet entourée d'autres structures radiales. Les autres coraux ne possèdent pas cette caractéristique morphologique singulière.

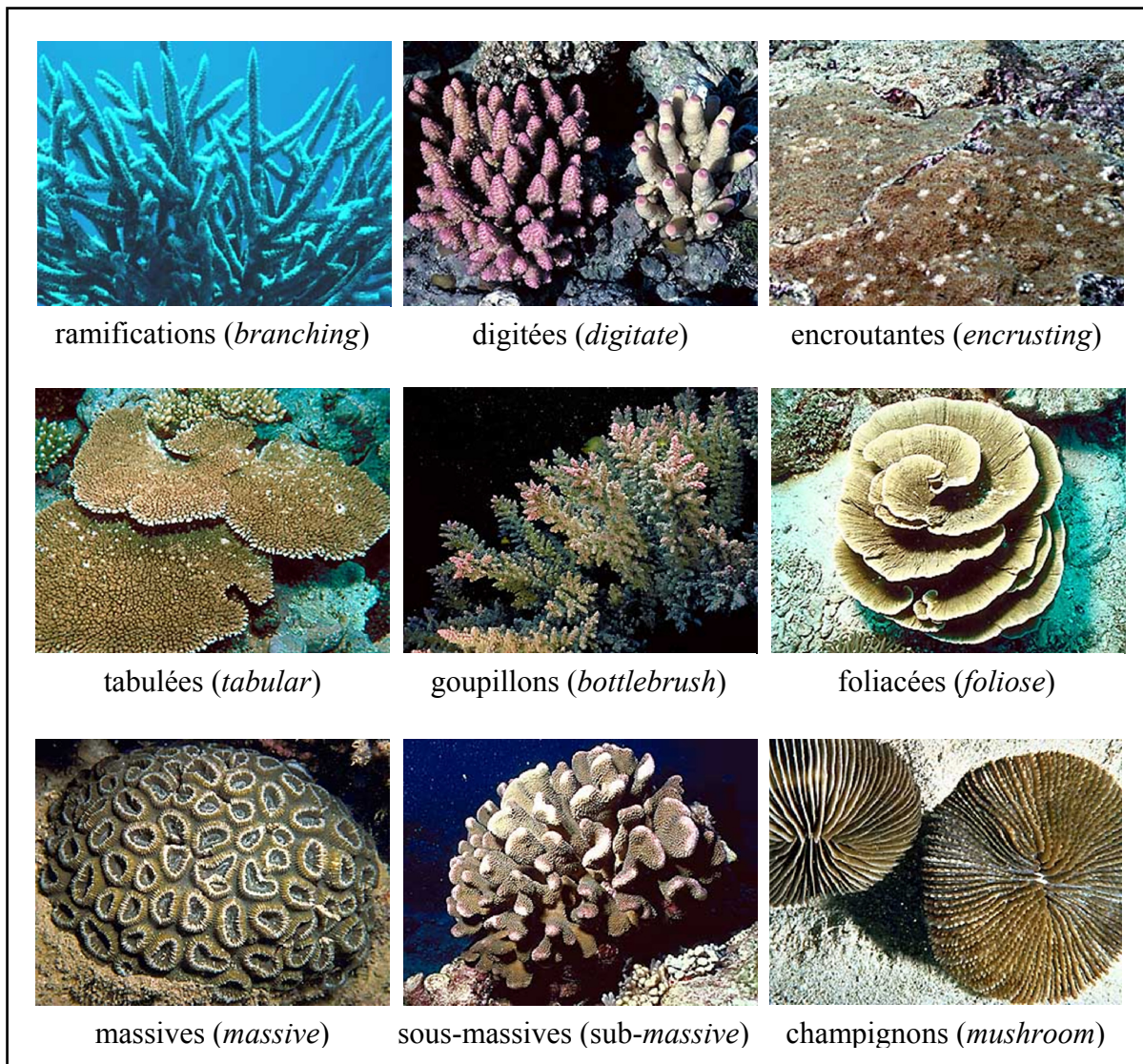


Figure 2.1 Exemples de formes benthiques communes.  
Adaptée de Licuanan (2001)

Tout compte fait, certaines formes benthiques n'ont pas encore été présentées. Les prochains paragraphes présenteront rapidement la description de ces formes benthiques particulières. En premier lieu, certaines explications sont nécessaires pour bien comprendre la nature des divers types de substrats marins. Les fonds marins étudiés sont ségrégués en cinq catégories :

- les coraux morts (*dead coral*), évidents à localiser par leur couleur blanche ou faiblement saturée. On parle ici d'un corail récemment mort, donc non recouvert d'algues;
- les coraux morts couverts d'algues (*dead coral with algae*), qui sont encore entiers, mais recouverts d'algues fines (*turf algae*);
- les moellons (*rubble*), en apparence de petits cailloux qui sont en fait des fragments de coraux morts non consolidés dans la roche;
- le sable (*sand*), sédiment facile à identifier;
- la vase (*silt*), sédiment très fin et qui a tendance à créer un brouillard, lorsqu'agité.

De plus, sur le substrat marin, on retrouve également plusieurs algues qui doivent être identifiées, le cas échéant. On dénote plusieurs identificateurs d'algues dont :

- les algues macroscopiques (*macro algae*), dont les individus (ou les « feuilles ») sont de taille supérieure à 5 cm;
- les algues coralliennes (*coralline algae*), qui se retrouvent généralement sur des coraux vivants, mais dans ce cas-ci, se sont déposées sur le substrat marin;
- les algues fines (*turf algae*) qui contrairement aux algues macroscopiques ont des individus de très petite taille ou quasi indiscernables;
- les assemblages (*algal assemblage*), sont des algues qui ne peuvent être identifiées avec plus de détails et qui ne font pas partie des trois dernières catégories.

Enfin deux dernières catégories de codes benthiques complètent le portrait. Premièrement, un groupe permet de classer les formes inconnues ou inhabituelles :

- les coraux mous (*soft coral*), qui sont faciles à identifier puisque dénués de structure;
- les éponges de mer (*sponge*);
- les autres objets (*other*), qui concernent tous les autres objets marins qui peuvent se retrouver dans les photographies, les poisons par exemple;
- les objets inconnus (*unknown*), lorsque la photographie est trop sombre ou trop hors foyer pour faire l'identification.

Finalement, trois genres sont identifiés plus précisément puisqu'ils sont populaires en aquariophilie et identifiables relativement facilement pour les initiés: les *heliopora*, les *millepora* (vénéneux d'où son surnom *fire coral*) et les *halimeda*. Cette identification nécessitant une connaissance préalable de ces espèces, nous n'entrerons pas dans les détails.

## 2.2 L'acquisition des images et les prétraitements

L'acquisition des images se fait selon un protocole qui permet une représentation uniforme de l'espace et constante dans le temps. Le secteur étudié est découpé en transects, soit en une multitude de lignes droites qui sont étudiées, année après année, afin d'effectuer le suivi. Le protocole pour mettre en place un échantillonnage par transects est décrit en détail par Page *et al.* (2001) et par Hodgson *et al.* (2004).

Pour cette étude, les données dont nous disposons sont des photographies numériques acquises par un plongeur se déplaçant en ligne droite au dessus du transect étudié. Tout au long du processus, le plongeur en question doit simplement contrôler la distance ( $\pm 1$  m) et perpendicularité de la caméra par rapport au benthos. Le déplacement en ligne droite est quant à lui contrôlé par le pilote de l'embarcation motorisée qui traine, avec un câble submergé, le plongeur dans la bonne direction. Étant hors de l'eau, le pilote de l'embarcation est en mesure de bien contrôler par GPS la rectitude de la trajectoire. De plus, un ruban à mesurer est disposé le long du transect et visible sur chaque image. Ce ruban assure également la rectitude de la trajectoire et sert de référence de position et d'échelle. Une photo est prise à tous les mètres environ. Diverses techniques sont maintenant utilisées pour uniformiser davantage la hauteur et la perpendicularité de la caméra, comme le montre la Figure 2.2. Toutefois, nos données n'ont pas été acquises avec ces nouveaux protocoles, les photographies ont été prises à main levée.

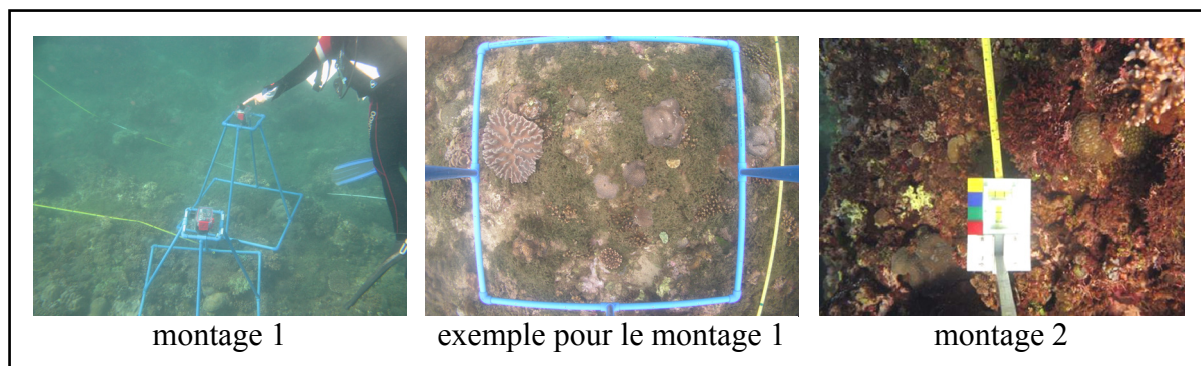


Figure 2.2 Divers montages pour la photographie sous-marine.

Le montage 1 de la Figure 2.2 assure la perpendicularité ainsi que la distance de prise de vue. L'autre montage assure simplement une distance constante et permet au plongeur d'assurer la perpendicularité via deux niveaux à bulle. (Ce qui peut être inefficace lorsque le substrat est très accidenté.) Toutefois, le montage 2 permet l'ajustement continu de la balance des couleurs avec sa référence blanche.

Malgré que notre banque d'images ait été prise à main levée, la balance des couleurs a été effectuée manuellement à chaque début de transect pour assurer la fidélité des couleurs qui est grandement affectée par l'eau. Pour ce faire, une cible blanche et quelques pastilles de couleur ont été utilisées de façon très similaire à l'image du montage 2 de la Figure 2.2. Le reste des paramètres de l'exposition des images (vitesse d'obturateur, ouverture de lentille et sensibilité ISO) ont été ajustés automatiquement par l'appareil pour assurer une exposition correcte. Finalement, des appareils compacts ont été utilisés, dans des caissons étanches. Ces appareils fournissent des images à une résolution spatiale de cinq mégapixels (rappelons que celles-ci ont été acquises en 2006).

### 2.3 La base de données

La base de données des Philippines est formée de trois entités distinctes. La première d'entre elles est la banque d'images originales, soit les images telles qu'acquises par les plongeurs (voir Figure 2.3, image originale). Suite à l'acquisition, un système automatisé superpose dix

points d'échantillonnage sur chaque image, sous forme de marqueurs cruciformes rouges (voir Figure 2.3, image avec marqueurs). Ces points d'échantillonnage sont plus tard référencés par des numéros de 1 à 10. Les numéros d'identification sont attribués de gauche à droite et de haut en bas. Enfin, l'objet se trouvant sous chaque marqueur doit être attribué à un des codes benthiques de la section 2.1 par un expert en la matière. L'identification de chaque marqueur est consignée dans la troisième entité de données, la table de référence (voir Figure 2.3, table de référence).

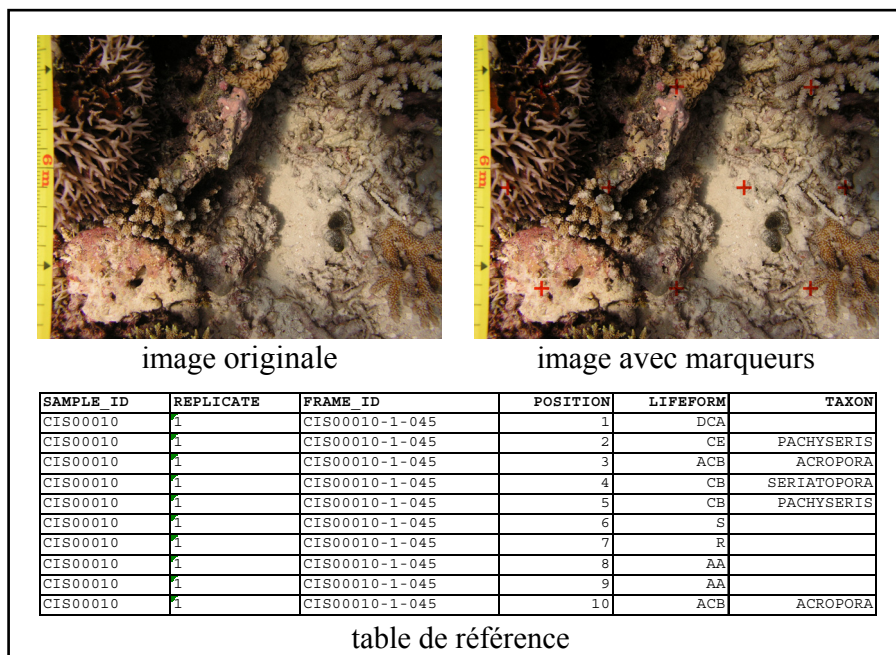


Figure 2.3 Les entités de la base de données des Philippines.

L'échantillonnage du benthos dont nous disposons est composé de 2 395 images en provenance de 24 transects différents. Considérant les dix marqueurs d'échantillonnage par image, nous disposerons potentiellement de 23 950 objets marins pour tenter l'automatisation du procédé d'identification. Ces données sont réparties de façon très hétérogène parmi les 28 classes (ou codes benthiques) énumérées à la section 2.1. Cette répartition sera détaillée plus tard, dans les prochaines sections. Toutefois, quelques codes incongrus se sont glissés parmi les données. En général, ces codes sont très peu utilisés. Ils ont donc été éliminés ou reclassés au cours des expérimentations. Le Tableau 2.2 liste ces codes et leur description détaillée.



Tableau 2.2 Les codes benthiques supplémentaires

Code benthique	Description	Utilisation des spécimens
WA	Water	Spécimens impossibles à classer dû à la grande distance de la photographie par rapport au substrat
DDD	Non-data points	Aucun spécimen visible sur ces photographies qui ont été prises pour montrer le paysage marin
SAMEPICWITH75	Duplicata	Erreur, photographie dupliquée
NOTSCORED	Not scored	Erreur, photographie sans marqueurs
UNK	Unknown	Spécimen inconnu, devrait plutôt être identifié par IND
GO	Nom inconnu	Code ou groupe benthique non explicité par le protocole, très peu utilisé

## 2.4 La segmentation

La structure particulière de la base de données, qui identifie la nature de dix points d'échantillonnage sur chaque image, permet d'utiliser un système de segmentation très basique. Aucune procédure complexe, inspirée de la vision artificielle (VA), n'est nécessaire dans ce contexte pour extraire la zone jugée pertinente à l'analyse sous chaque marqueur. Nous avons plutôt choisi d'utiliser une simple zone rectangulaire de 96 par 72 pixels, centrée sur le marqueur. Cette taille de zone semble un bon compromis puisque sa taille est suffisante pour l'extraction de descripteurs et est également suffisamment petite pour caractériser localement la zone identifiée par le marqueur. De plus, elle permet de diviser l'image en un nombre entier de sous-images. Nous avons choisi d'utiliser ce point de départ pour s'approprier toute la complexité de la problématique, malgré qu'une étude précédente n'ait pas été concluante sur la taille à optimiser (voir la section 1.2.2).

## 2.5 Les expérimentations sur l'approche de description par formes benthiques

Maintenant que les contextes scientifique et logiciel sont connus, cette section présentera les diverses expériences que nous avons effectuées sur la base de données des Philippines. Ces expérimentations ont toutes pour but d'isoler le ou les éléments problématiques qui causent les mauvaises performances observées au cours des travaux précédents (voir la section 1.2, page 16). Les expériences seront présentées en ordre chronologique, tel que résumé par le Tableau 0.1 de la page 7. Notez qu'à partir de maintenant, la terminologie relative à la théorie de la reconnaissance de formes (RDF) sera utilisée. Nous ne parlerons plus d'*identification*, mais plutôt de *classification*, ni de *code benthique*, mais plutôt de *classe*.

### 2.5.1 L'optimisation de la représentation des données

La première hypothèse testée pour résoudre le problème de classification est la suivante : *le fort déséquilibre dans la représentation des classes est cause des performances médiocres*. En effet, une des observations que l'on peut tirer des matrices de confusion résultantes de nos tests préliminaires est que la classe AA (*algal assemblage*) qui est très fortement représentée dans la base de données, par rapport à toutes les autres classes, devient la spécialité de notre classificateur. En fait, puisque plus de 40 % des données disponibles appartiennent à la classe AA, le processus d'optimisation à la base de l'apprentissage automatisé tend à classer toutes les données disponibles dans cette classe. Il obtient ainsi très facilement un optimum local à un taux de bonne classification de l'ordre de 40 %. De plus, cette représentation fortement hétérogène implique un nombre très faible d'échantillons pour certaines classes. Ce faible nombre d'individus semble lié aux taux de classification très médiocres observés pour les classes concernées.

#### 2.5.1.1 Méthodologie

La première tentative, présentée par Bouchard (2008a), est de rééchantillonner la base de données pour obtenir un classificateur qui prend équitablement en considération chacune des

classes. Cette expérimentation tente également de valider l'impact de la faible représentation de certains codes benthiques sur leur taux de classification. Pour ce faire, des échantillons de 96x72 pixels ont été prélevés sous chacun des dix points d'échantillonnage des images. Puis, les descripteurs proposés par Levasseur (2008) ont été extraits de ces échantillons (voir la section 1.2.1). Notez que les descripteurs ont été normalisés entre zéro et un pour les expérimentations. Ainsi, aucun d'entre eux n'a préséance sur les autres par sa plage étendue de valeurs. À partir de la base de données au format ARFF, qui découle de l'extraction de caractéristiques, neuf échantillonnages différents ont été réalisés (voir le Tableau 2.3).

Tableau 2.3 Les échantillonnages testés par l'expérimentation

<b>Échantillonnage</b>	<b>Nb. de classes</b>	<b>Nombre d'échantillons</b>	<b>Ratio</b>
La base de données originale	30	21 847	100,0 %
Les classes avec 350 échantillons et plus	14	20 297	92,9 %
Exactement 350 échantillons par classe <sup>a</sup>	14	4 900	22,4 %
Classes avec 1000 échantillons et plus	7	16 864	77,2 %
Exactement 730 échantillons par classe <sup>b</sup>	30	21 900	100,2 %
Base originale sans données aberrantes globales <sup>c</sup>	30	19 541	89,4 %
Base originale sans données aberrantes de classe <sup>d</sup>	30	19 707	90,2 %
Exactement 660 échantillons par classe <sup>e</sup>	30	19 800	90,6 %
Classes avec 900 échantillons et plus, sans aberrances	7	4 620	21,1 %

<sup>a</sup> Seulement les classes avec 350 échantillons. Pige aléatoire sans remise.

<sup>b</sup> Sujet à une pige aléatoire sans remise pour les classes avec plus de 730 échantillons. Une pige avec remise a été effectuée pour les autres classes en plus de l'ajout d'une variance aléatoire de  $1,5\sigma$ . (L'écart-type  $\sigma$  est calculé sur la distribution initiale de la classe.)

<sup>c</sup> Une donnée est considérée aberrante lorsqu'elle se situe à plus de trois écarts interquartiles des quartiles extérieurs. (Les quartiles sont calculés sur la distribution entière des données.)

<sup>d</sup> Idem à *b*. Toutefois, les quartiles sont calculés individuellement pour chaque classe.

<sup>e</sup> Idem à *a*. Toutefois, la distribution exclue les données aberrantes détectées en *b*.

Enfin, cinq paradigmes de classification différents soient : l'algorithme naïf de Bayes (NB), les plus proches voisins (kPPV), l'arbre de classification C4.5 (C4.5), le séparateur à vaste marge (SVM) et le perceptron multicouche (PM) ont été utilisés pour tenter l'automatisation. (Voir la section 1.3.5 pour une description complète.) Le Tableau 2.4 décrit sommairement les métaparamètres utilisés au cours de l'expérience pour les algorithmes. Ces paramètres ont été déterminés à priori selon les recommandations générales pour chacun des paradigmes de classification, tels que proposés par Hall *et al.* (2009). Les valeurs n'ont pas été optimisées puisque ce n'est pas le but précis de cette expérimentation.

Tableau 2.4 Les métaparamètres généralement recommandés

Algorithme	Paramètre	Valeur
NB	aucun paramètre	non applicable
kPPV	nombre de voisins (k)	3
	distance	linéaire euclidienne
C4.5	niveau de confiance (c)	0,25
	émondage	aucun
SVM	facteur de complexité (C)	1,0
	epsilon ( $\epsilon$ )	$1,0 \times 10^{-12}$
	noyau	polynomial de degré 1
	tolérance	0,0010
PM	nombre de couches cachées	1
	nombres de neurones sur couche cachée	(nb. attributs + nb. classes)/2
	taux d'apprentissage	0,3
	momentum	0,2
	nombre d'époques	500

### 2.5.1.2 Résultats expérimentaux

Au cours de l'apprentissage avec ces algorithmes, 66 % des imagerie de la base de données (images de 96x72 pixels centrées sur les marqueurs d'échantillonnages décrits en 2.3) ont été utilisées. Finalement, les tests ont été exécutés sur le tiers restant. Tout ce processus, incluant l'entraînement et la généralisation, a été répété six fois, en variant aléatoirement la portion d'apprentissage et de test de la base de données, pour vérifier la stabilité des résultats. Le Tableau 2.5 présente les résultats moyens obtenus au cours de cette expérience.

Tableau 2.5 Taux de classification moyens reliés aux différents échantillonnages

<b>Expérience</b>	<b>NB</b>	<b>kPPV</b>	<b>C4.5</b>	<b>SVM</b>	<b>PM</b>
La base de données originale	9,6	34,9	26,4	39,1	40,6
Les classes avec 350 échantillons et plus	17,4	31,6	29,0	42,2	43,0
Exactement 350 échantillons par classe	20,8	19,7	18,0	26,8	25,2
Classes avec 1000 échantillons et plus	27,4	44,7	36,2	47,8	49,6
Exactement 730 échantillons par classe	47,5	20,6	34,2	33,4	33,9
Base originale sans données aberrantes globales *	12,4	34,2	24,5	37,7	39,3
Base originale sans données aberrantes de classe	14,9	37,2	28,9	41,6	42,5
Exactement 660 échantillons par classe	49,8	25,5	34,8	39,2	38,6
Classes avec 900 échantillons et plus, sans aberrantes	29,8	46,5	37,6	49,7	51,3

\* Un échantillon est considéré aberrant si au moins une de ses caractéristiques se situe à plus de trois écarts interquartiles de la médiane de sa distribution.

*Note : Les taux de classification sont présentés en pourcentage de classification correcte. L'importance de ces résultats et l'écart-type étant tous deux très faibles, aucune analyse statistique n'a été complétée.*

### 2.5.1.3 Discussion

Selon les résultats obtenus, notre hypothèse ne tient plus. Il est effectivement possible d'améliorer les performances en rééchantillonnant la base de données. Toutefois, l'échantillonnage disproportionné ne semble pas être la principale cause des faibles performances. Malgré tout, au cours de cette première tentative plusieurs découvertes importantes ont été faites. En premier lieu, le nombre d'échantillons réellement disponibles est 21 847, ce qui est différent des 23 950 échantillons envisagés par notre description de la base de données en 2.3. Cet écart est dû à la divergence entre la table de référence et la banque d'images. En fait, cette discordance est d'ordre majeur et est liée à la grande quantité de données aberrantes trouvées. La prochaine expérience s'intéresse à ce problème.

### 2.5.2 La restructuration de la base de données des Philippines

Suite à la découverte d'une grande quantité de données aberrantes dans les images en 2.5.1, une étude plus poussée sur le sujet a été présentée par Bouchard (2008b). Au cours de ce travail, l'hypothèse maintenue est qu'*une incohérence majeure entre les diverses entités de données est la principale cause des performances sous-optimales observées*. En effet, il est apparu que « les noms de fichiers utilisés pour les images avec marqueurs et sans marqueurs ne correspondent pas toujours ». Ce problème influence également le lien entre la table de référence (qui indique les codes benthiques des échantillons) et la banque d'images. En d'autres termes, les échantillons prélevés ne sont pas attribués à leurs classes véritables.

La banque d'images originales (sans marqueurs) compte 2 635 clichés. La banque d'images avec marqueurs compte quant à elle 2 395 clichés. La différence (240 images) entre ces deux entités est explicable par deux types de photographies : les photos pour la balance des blancs et certaines photographies prises hors de tout contexte scientifique (diverses scènes marines). Les diverses scènes marines, non pertinentes à l'étude de la biodiversité, ont été prises en compte dans la numérotation des images originales, mais *n'ont pas été prises en*

*considération lors de la numérotation des images avec marqueurs.* Cette discordance a créé une confusion majeure, suite au décalage des indices des photographies numérotées.

### **2.5.2.1 Méthodologie et résultats expérimentaux**

Pour résoudre cette problématique, nous avons implémenté un algorithme pour comparer les images, avec et sans marqueurs, qui établit une table de correspondance fiable entre les trois entités de données. Enfin, même du côté de la table de référence certaines images étaient manquantes. Une grande quantité d'identifications ont été validées manuellement. Suite à tout ce travail, une grande quantité d'échantillons ont été corrigés; le Tableau 2.7 de la page 83 en indique les proportions. Le Tableau 2.6 de la page 82 indique quant à lui, la véritable répartition des données parmi les classes. Cette répartition des classes n'a pas été présentée dans la section traitant de la base de données en 2.3 pour éviter toute confusion.

Suite à la correction de la classification d'une énorme quantité d'images nous avons dû reprendre les expériences. Dans un premier temps, l'algorithme d'extraction des caractéristiques fourni par Levasseur (2008) a été optimisé et validé. De plus, une référence à l'image d'origine a été ajoutée à la base de données en format ARFF pour permettre un suivi. Enfin, diverses expériences ont été complétées pour valider certains échantillonnages. Ces expériences sont résumées au Tableau 2.8 de la page 84. Ces expériences tentent d'améliorer les résultats de classification en imbriquant diverses modifications au système. Notez que le même protocole expérimental qu'en 2.5.1 a été utilisé. Soit, une répartition deux tiers / un tiers pour l'apprentissage et les tests. Nous avons aussi utilisé les cinq mêmes classificateurs et leurs métaparamètres. Le premier test exécuté est évidemment l'apprentissage et la classification des 19 530 échantillons corrigés (ce total est expliqué un peu plus loin). Ensuite, au cours d'autres expériences comprenant l'apprentissage et la classification, des classes ont été retirées pour simplifier le problème et prioriser les classes (codes benthiques) d'intérêt. Tout d'abord, les spécimens *non identifiés* ont été retirés : soit les classes DDD, IND, NOTSCORED, UNK et SAMEPICWITH75. Pour épurer davantage les données, d'autres classes ont aussi été retirées soit : les *non-vivants* OT, R, S, SI et WA.

Tableau 2.6 Répartition des échantillons parmi les classes de la base de données

Catégorie	Compte	Groupe	Compte	Type	Description <sup>a</sup>	Compte	Ratio
Corail3	13950 71,5 %	Abiotique	387 2,0 %	DC	Dead coral	98	0,5 %
				DCA	Algae cov. cor.	289	1,5 %
		Acropora	1954 10,0 %	ACB	Branching	1478	7,6 %
				ACD	Digitate	15	0,1 %
				ACE	Encrusting	14	0,1 %
				ACS	Submassive	68	0,3 %
				ACT	Tabulate	370	1,9 %
				ACX	Bottlebrush	9	0,0 %
				Algues	7141 36,6 %	AA	Algal assemb.
		CA	C. algae			407	2,1 %
		HA	Halimeda			73	0,4 %
		MA	Macro algae			340	1,7 %
		TA	Turf algae			28	0,1 %
		Non-acropora	2750 14,1 %	CB	Branching	975	5,0 %
				CE	Encrusting	1152	5,9 %
				CF	Foliaceous	163	0,8 %
				CM	Massive	305	1,6 %
				CMR	Mushr. coral	83	0,4 %
				CS	Submassive	72	0,4 %
		Autres	1718 8,8 %	CHL	Heliopora	67	0,3 %
CME	Millepora			212	1,1 %		
SC	Soft coral			971	5,0 %		
SP	Sponge			468	2,4 %		
Non-corail	5560 28,5 %	Abiotique	3151 16,2 %	R	Rubble	2003	10,3 %
				S	Sand	995	5,1 %
				SI	Silt	1	0,0 %
				WA	Water	152	0,8 %
		Autres	2409 12,3 %	DDD	Non data pt.	1146	5,9 %
				GO	(inconnue)	47	0,2 %
				IND	Inderterminate	696	3,6 %
				OT	Animal	455	2,3 %
				UNK	Unknown	65	0,3 %
				Total			

<sup>a</sup> Tout comme au Tableau 2.1 (p. 69) la description a été volontairement laissée en anglais.

<sup>b</sup> Les classes NOTSCORED et SAMEPICWITH75 ont été ignorées ce qui affecte le total.



Tableau 2.7 Statistiques sur la restructuration de la base de données des Philippines

<b>Statistique</b>	<b>Nombre</b>	<b>Ratio</b>
Échantillons répertoriés dans la table de référence	26 790	100 %
Échantillons éliminés : image originale introuvable	2 500	9 %
Échantillons éliminés : image avec marqueurs introuvable	4 020	15 %
Échantillons éliminés : aucune des deux images n'est disponible	730	3 %
Échantillons éliminés : image corrompue sans marqueur	10	~ 0 %
Échantillons véritablement disponibles pour l'analyse	19 530	73 %
Échantillons corrigés (réalignement des indices des images)	7 650	<sup>a</sup> 39 %

<sup>a</sup> Ratio calculé par rapport à la quantité d'échantillons véritablement disponibles.

Puisque « le traitement des données est effectué manuellement et par plusieurs personnes lors de la classification des spécimens », un problème subsiste. « Certains spécialistes ont classifié le spécimen le plus près du marqueur lorsque celui-ci ne pointait pas sur un sujet intéressant. Ils auraient plutôt dû rigoureusement indiquer ce qui se situe sous le marqueur » (Bouchard, 2008b). Pour cette raison, une analyse des données aberrantes a été complétée. Toutefois, comme les résultats n'étaient toujours pas satisfaisants à cette étape (voir le Tableau 2.8), nous avons procédé au rééchantillonnage (pige avec remise) des données pour assurer un nombre équivalent d'échantillons par classe (retour à l'hypothèse de 2.5.1).

Malgré tout ce travail, le problème des faibles taux de classification n'a pas été résolu. Un léger travail sur les caractéristiques a donc été entamé. La projection des données dans l'espace simplifié de l'analyse par composantes principales (ACP) couvrant 95 % de la variance a été testée. (L'ACP et les transformations subséquentes ont été effectuées à partir des données en excluant les spécimens non identifiés, non vivants et aberrants.) Enfin, même le processus d'entraînement des classificateurs a été modifié. Un entraînement avec validation croisée à dix partitions a été testé.

Tableau 2.8 Résultats obtenus sur la base de données des Philippines restructurée

Échantillonnage	NB	kPPV	C4.5	SVM	PM
Taux de classification de l'étude précédente	9,6	34,9	26,4	39,1	40,6
Taux de classification des nouvelles données corrigées	11,2	37,5	29,5	38,6	41,6
Taux de classification sans les spécimens non identifiés	9,6	37,2	28,3	37,9	41,1
Taux de classification sans les spécimens non vivants	9,6	44,0	32,9	45,1	47,2
Taux de classification sans les données aberrantes	12,5	47,1	35,8	48,2	51,2
Taux de classification avec rééchantillonnage	5,9	20,0	21,0	19,7	36,0
Taux de classification avec ACP	44,2	46,4	35,6	47,1	51,1
Taux de classification avec ACP et rééchantillonnage	9,0	18,4	21,5	14,2	14,2
Taux de classification avec validation croisée	13,0	47,0	36,4	47,6	50,7
Taux de classification avec validation croisée et ACP	43,6	46,3	37,0	46,5	50,2

### 2.5.2.2 Discussion

Enfin, après les divers essais, on remarque que cette deuxième hypothèse est également fautive. Bien que près de 40 % des données utilisables aient été réétiquetées correctement, aucun impact marqué sur les performances n'a été observé. Ceci met en évidence une chose : les algorithmes de RDF sont tolérants au bruit. De plus, avec l'expérimentation de l'ACP, nous constatons que malgré le vaste ensemble de caractéristiques utilisées nous ne sommes probablement pas victimes de la *malédiction de la dimensionnalité* telle que décrite par Richard E. Bellman (Bellman, 1961). En d'autres termes la qualité de la solution optimale ne semble pas influencée par la grandeur de l'espace de recherche.

### 2.5.3 L'implémentation d'un système étagé avec mécanisme de rejet

Il est évident qu'un autre problème subsiste toujours au sujet de ce système ou des données. Les deux prochains travaux présenteront diverses modifications à l'architecture du système de classification. L'hypothèse posée stipule donc que *l'architecture du système de*

*classification est la cause des performances médiocres.* À cette étape, il aurait été préférable de vérifier si l'ensemble des caractéristiques extraites est adéquat pour le problème de reconnaissance des coraux.

### **2.5.3.1 Méthodologie**

Au cours de cette investigation, deux nouvelles méthodes ont été utilisées pour améliorer les performances : l'implémentation d'un classificateur étagé avec un module *détecteur* accompagné d'un module *classificateur* et l'ajout d'un mécanisme de *rejet*. Cette nouvelle architecture permet de découpler la tâche de détection du corail, ce qui simplifie grandement la tâche de classification. Toutes les classes de la catégorie non corail sont immédiatement mises de côté alors que les coraux poursuivent leur chemin vers le classificateur. L'engin de classification se spécialise donc sur le corail et non sur les spécimens de moindre intérêt (les non-coraux, voir Tableau 2.6).

Le mécanisme de rejet permet quant à lui de fixer le taux de confiance sur la décision du système de classification. Les spécimens seront classifiés à un taux de confiance de l'ordre des 80 %, ce qui est comparable à l'efficacité de la classification manuelle. Et finalement, les cas difficiles, où le taux de confiance sur la décision est inférieur à un certain seuil, sont simplement rejetés par ce mécanisme. Ces spécimens rejetés devront être classifiés manuellement par un expert. Notez que ce mécanisme de rejet peut être implémenté aux deux niveaux du système de classification en cascade, soit au niveau du *détecteur* et du *classificateur*. La Figure 2.4 montre le système en cascade avec rejet.

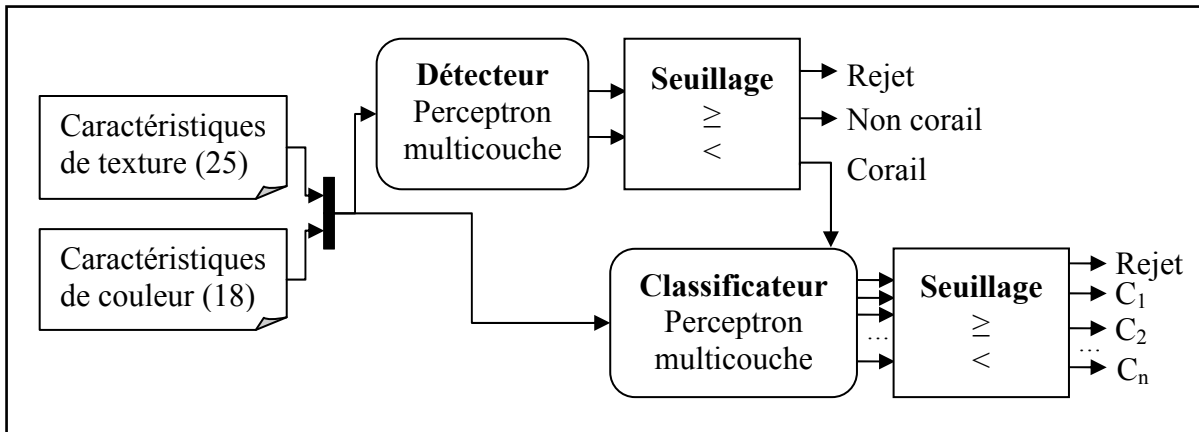


Figure 2.4 Le système étagé avec mécanisme de rejet.  
Adaptée de Bouchard (2009c), idée originale de Landgrebe *et al.* (2005)

Cette fois, seul le perceptron multicouche a été utilisé, puisque jusqu'à maintenant c'est l'algorithme le plus performant. Une configuration similaire aux études précédentes des métaparamètres a été utilisée (voir Tableau 2.4 p. 78). En revanche, le PM n'a pas été entraîné sur 500 époques, mais plutôt sur le nombre minimal d'époques nécessaires. Nous évitons ainsi le surapprentissage. Pour y arriver, 15 % de la base d'apprentissage a été réservée à la validation du point d'arrêt. Lorsque 20 époques successives entraînent l'augmentation de l'erreur sur cette portion réservée pour la validation, l'entraînement est arrêté. De plus, 20 % des données ont été mises de côté pour les tests cette fois, contrairement au 33 % utilisées précédemment. L'entraînement du PM a été effectué au sein d'un processus de validation croisée à 10 partitions et tout sous-échantillonnage de la base de données a été effectué de façon stratifiée. La stratification assure la conservation des proportions entre les classes, telle qu'elle est dans la base de données initiale. Finalement, les données corrigées en 2.5.2 ont été utilisées.

Suite à l'entraînement, le mécanisme de rejet entre en action. Deux méthodes différentes ont été testées pour établir les seuils de rejet (Chow, 1970; Fumera, Roli et Giacinto, 2000). Le seuillage de Chow repose sur une valeur de seuil unique pour toutes les classes de la base de données, tandis que du côté de Fumera un seuil adapté à chaque classe est proposé. Référez-vous à la section 1.3.6.2 pour de plus amples renseignements.

### 2.5.3.2 Résultats expérimentaux

Le Tableau 2.9 indique que ces modifications au système sont une avenue très intéressante.

Tableau 2.9 Résultats obtenus sur Philippines avec un système étagé

Système	Taux de classification	Taux de rejet
De base (tiré du Tableau 2.8)	41,6 %	0 %
Détecteur seulement	78,6 %	0 %
Classificateur de coraux seulement	42,7 %	0 %
Système étagé sans rejet	50,3 %	0 %
Système étagé avec rejet de Chow	88,47 %	57,1 %
Système étagé avec rejet de Fumera	82,60 %	47,8 %

### 2.5.3.3 Discussion

En réponse à l'hypothèse, l'architecture du système est un facteur qui peut en effet influencer positivement les performances. Toutefois, pour obtenir une performance acceptable avec rejet, une grande quantité de données doivent encore être traitées manuellement. Notez que pour les résultats présentés, un mécanisme de rejet a été implanté à chaque étage. Le traitement manuel exigera donc à l'utilisateur d'identifier si un spécimen rejeté est un corail ou non et/ou quel est le code benthique associé au corail.

Enfin, la théorie de Fumera affirme que son seuillage est toujours plus performant que celui de Chow. Toutefois, la tâche d'ajustement des seuils est très complexe. Explorer tout l'espace des paramètres pour en faire l'optimisation s'avère beaucoup trop fastidieux. Un algorithme d'optimisation simple a été utilisé (voir Figure 1.12), voilà pourquoi nos résultats ne concordent pas exactement avec ceux de Fumera. Un algorithme holistique tel

l'optimisation par algorithmes génétiques (Holland, 1992) ou par essaim de particules (Eberhart et Kennedy, 1995) serait préférable.

#### **2.5.4 L'implémentation d'un système d'identification par détecteurs**

Puisque les performances peuvent être améliorées par un système étagé, ce travail s'intéresse davantage aux différentes configurations possibles (Bouchard, 2009b). L'hypothèse posée ici est un peu plus précise que la précédente et stipule que : *le découplage du système de détection et de classification cause une amélioration de performance*. En effet, les performances observées au cours de l'expérimentation précédente portent à croire qu'un système en cascade avec de multiples classificateurs permet la spécialisation sur des tâches précises, ce qui améliore les performances globales.

##### **2.5.4.1 Méthodologie**

Le système proposé découple donc davantage la détection de la classification. Les données sont d'abord présentées à un ensemble de cinq *détecteurs* chargés de détecter le groupe benthique de l'échantillon (voir la colonne « groupe » du Tableau 2.6). À chaque détecteur de groupe est associé un *classificateur* qui se spécialise dans la classification des classes de ce groupe. La Figure 2.5 présente un schéma de ce système complexe.

Le mécanisme de détection sélectionne le bon groupe benthique par un procédé de fusion de la décision. La règle choisie pour implémenter cette fusion détermine le groupe benthique le plus probable pour un échantillon en sélectionnant le score de détection le plus fort. Suite à une détection correcte du groupe de l'échantillon, il est acheminé au classificateur spécialisé sur ce groupe, qui se charge d'identifier le code benthique, parmi une liste réduite de codes.

Cette fois, la base de données a été divisée en deux partitions égales pour l'apprentissage et les tests. L'apprentissage a été effectué dans un processus de validation croisée à dix partitions ou une des partitions a été utilisée pour valider un des métaparamètres.

Contrairement aux expérimentations précédentes, le métaparamètre le plus important a été optimisé au cours de l'apprentissage. Tous les sous échantillonnages ont été séparés de façon stratifiée pour conserver la probabilité à priori des classes.

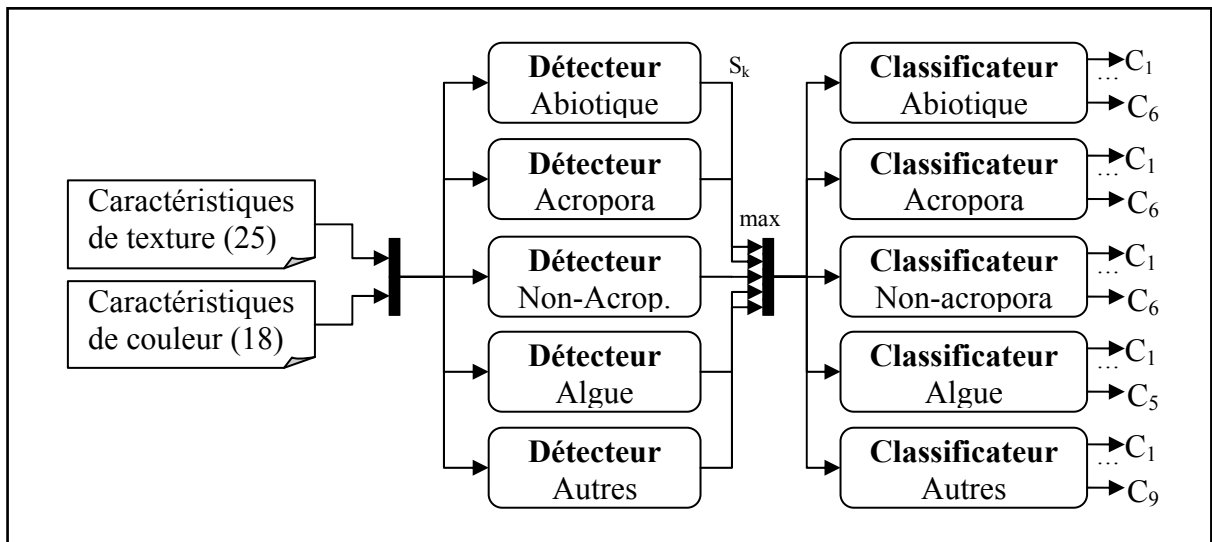


Figure 2.5 Le système d'identification par détecteurs.  
Adaptée de Bouchard (2009b)

Cette fois, deux classificateurs ont été mis à l'essai : le PM et le réseau de neurone à base radiale (RBR). Les métaparamètres utilisés sont décrits au Tableau 2.10. Pour chaque itération de la validation croisée, pour chaque groupe benthique et pour chaque valeur de métaparamètre possible, un détecteur a été entraîné sur la portion d'apprentissage (9 partitions sur 10) et testé sur la portion validation (1 partition sur 10) de la base de données d'apprentissage. Suite à ce processus, la meilleure paramétrisation de l'algorithme est sélectionnée en fonction de la performance obtenue en validation. Ensuite, les classificateurs de chaque groupe ont été entraînés de la même manière. À la fin d'une itération de la validation croisée, un système complet est entraîné et est mis à l'épreuve sur la base de test. Tout le processus d'entraînement, de validation et de test se solde donc en dix valeurs indépendantes de performance pour le système dans son ensemble.

Tableau 2.10 Méta-paramètres pour le système d'identification par détecteurs

Algorithme	Paramètre	Valeur
PM	nombre de couches cachées	1
	nombres de neurones sur couche cachée	(nb. attributs + nb. classes)/2
	taux d'apprentissage (optimisé sur une plage)	[0,1; 0,3; 0,5; 0,7]
	momentum	0,2
	nombre d'époques maximal	300
	partition de validation	13 %
	seuil de validation	20 époques
RBR	nombre maximal d'itérations	200
	écart type minimal pour les noyaux	0,1
	nombre de noyaux (optimisé sur une plage)	[33; 38; 43; 48]

#### 2.5.4.2 Résultats expérimentaux et discussion

Enfin, après l'expérimentation sur cette architecture de système, un taux moyen de classification de 34,96 % caractérisée par une variance de 0,04 % a été observé pour l'algorithme PM. Du côté du RBR, 26,55 % des échantillons ont été classifiés correctement (variance observée de 0,01 %). Ces résultats ne sont malheureusement pas très concluants. Considérés individuellement, les taux de classification de l'étage de détection et ceux de l'étage de classification sont pourtant acceptables (de l'ordre de 60 à 70 %). Il semble que le problème ait lieu au niveau de la fusion de l'information. Un travail plus poussé sur les méthodes de fusion pourrait être une piste de solution. Néanmoins, l'architecture de ce type de système étant très complexe, il est fort possible qu'un problème de surapprentissage soit une cause sous-jacente aux performances insatisfaisantes.



## 2.6 Discussion générale sur l'approche descriptive par formes benthiques

Malgré tous les efforts fournis, les conclusions des diverses expérimentations nous ont conduites à laisser de côté cette source de données pour le moment. Tel que mentionné en 2.1, une forme benthique est facilement identifiable par l'homme, et c'est d'ailleurs le but et tout le génie de l'approche. Avec ses connaissances à priori sur l'environnement, l'homme est capable de visualiser le spécimen dans l'espace et d'y attribuer l'une ou l'autre des étiquettes benthiques à partir d'une simple image. Ceci nous rappelle que le concept de forme benthique est indéniablement lié à la perception *tridimensionnelle* de l'objet. Dans ces conditions, tenter d'automatiser l'étiquetage à partir d'une image *bidimensionnelle* et un lot de descripteurs issus du plan image peut sembler contradictoire.

Par contre, selon les premières observations qualitatives effectuées sur diverses bases de données, il existe une corrélation importante entre la texture et la couleur d'un corail et son espèce. De plus, il existe une autre relation importante entre l'espèce et la forme de croissance (forme benthique) d'un spécimen. Le document de l'AIMS cité en 2.1 fournit à ce sujet une table partielle des associations possibles entre espèces et formes benthiques (Page *et al.*, 2001). Puisque ces deux relations existent, le système intelligent développé dans le cadre de ce projet devrait théoriquement être en mesure d'accomplir le travail.

Toutefois, une forme benthique peut être attribuée à un très grand nombre d'espèces distinctes, possédant des caractéristiques de texture et de couleur très variables. Inversement, on observe chez certaines espèces diverses formes de croissances, en fonction des conditions locales (température, courants marins, etc.). Ce phénomène complique le problème d'identification, lorsque l'on ne considère qu'uniquement l'information facilement accessible véhiculée par l'image numérique (la texture et la couleur). Malgré la complexité de la relation qui unit l'espèce et la forme benthique, elle est, comme bien des choses, de nature finie. Il est donc possible de la décrire entièrement, *si et seulement si* nous avons en notre possession suffisamment de données qui caractérisent la matérialisation des formes benthiques; assez de données pour bien caractériser leur relation avec l'espèce.

C'est donc la faible quantité de données par classes, combinée à la grande complexité de la relation entre les caractéristiques d'une image bidimensionnelle et les formes benthiques qui nous ont conduits à laisser de côté les données des Philippines. Finalement, divers éléments relevés suite à la rencontre de M. Licuanan et M. Landry en mars 2009 nous poussent également dans cette direction. Il semble qu'avec la disponibilité croissante des connaissances en taxinomie et des technologies liées à l'identification ADN, le suivi écologique par forme benthique soit devenu marginal. L'utilisation de la description taxinomique complète est dorénavant privilégiée par la communauté scientifique. En plus d'être plus précise, elle permet à tous de communiquer ses résultats dans un langage commun, qui s'étend à tous les domaines de la biologie.

Cependant, cette conclusion n'infirme pas notre hypothèse de travail, mais y ajoute une condition préalable: une identification visuelle macroscopique est suffisante pour départager la majorité des spécimens de corail, *considérant qu'assez de données sont disponibles pour caractériser la relation entre le stimulus visuel et l'identité du spécimen*. Il existe plusieurs façons d'accroître la quantité de données et rendre ce problème solvable. En retournant à l'origine des images (dans le « monde réel » selon la Figure 0.3 p. 6), une caméra stéréoscopique pourrait être utilisée pour permettre le calcul de descripteurs basés sur l'information tridimensionnelle. Plus encore, l'imagerie multispectrale pourrait également être employée. C'est connu, les coraux sont sujets à des comportements particuliers lorsqu'étudiés sous des longueurs d'onde hors du spectre du visible (Gleason, Reid et Voss, 2007). Enfin, peu importe la méthode, il est impératif de disposer d'assez d'images pour contrebalancer la complexité de discrimination des classes, qu'elle se manifeste par une grande variabilité intraclasse ou une grande confusion interclasse. Malgré tout, les recherches effectuées sur ces données ont permis de développer une architecture de détection et de classification de corail qui peut être utilisée dans un contexte différent. Ce sera le point de départ de notre prochaine expérimentation qui traite de l'identification la plus granulaire possible : l'identification par description taxinomique complète. *Prenez note que puisque les expérimentations sur ces données (des Philippines) n'ont pas été concluantes, nous n'avons pas validé statistiquement leurs résultats à grande échelle.*

## CHAPITRE 3

### APPROCHE PAR DESCRIPTION TAXINOMIQUE COMPLÈTE

Suite aux conclusions tirées en 2.6 sur la description par formes benthiques, il n'y avait d'autre choix que de reprendre l'analyse sur une tout autre base de données dotée d'un niveau de classification atomique. Suite à de plus d'amples recherches, il est apparu que très peu de ressources photographiques fiables identifient les coraux au niveau taxinomique. Seules les ressources encyclopédiques fournissent, en faible nombre, des images correctement identifiées. Deux de ces ressources ont attiré notre attention : *CoralPedia* (Sheppard, 2007) qui dresse un portrait des coraux des caraïbes et *Coral ID* (Veron et Stafford-Smith, 2002) qui traite des spécimens scléactiniens du monde entier. Pour cette série d'expérimentations, nous avons choisi d'utiliser *Coral ID* pour les raisons suivantes. Premièrement, *Coral ID* propose quelques images pour chaque espèce présente aux Philippines, tel que documenté par Licuanan (2008). De plus, contrairement à *CoralPedia*, *Coral ID* est associé à une publication (Veron, 2000) très reconnue dans le domaine de la taxinomie des coraux.

L'hypothèse que nous vérifierons ici est formulée en fonction de l'expérience acquise avec la représentation par formes benthiques : *l'entraînement d'un système intelligent d'identification de corail est possible à l'aide d'une banque d'images classifiées à un niveau atomique*. Tel que démontré lors de la première partie de notre expérimentation (section 3.1), l'atomicité des classes est par définition respectée par la description taxinomique complète. De plus, nous considérons toujours l'importante condition ajoutée à l'hypothèse générale en 2.6 : *assez de données doivent être disponibles pour caractériser la relation entre le stimulus visuel et l'identité du spécimen*. Si cette condition est respectée, notre système devrait être en mesure de bien généraliser sur l'espèce à partir de l'information de l'image. Toutefois, nous posons également l'hypothèse que *puisque les variabilités inter-classe et intra-classe sont maintenant réduites par l'identification taxinomique, la quantité de données nécessaire pourra être moins importante pour assurer le bon fonctionnement du système de*

*classification*. Pour valider cet ensemble d'hypothèses, nous évaluerons l'impact du nombre de classes considérées (en utilisant toute l'information disponible de la base de données) et inversement, nous évaluerons l'impact de la quantité d'information utilisée, considérant un nombre fixe d'espèces à départager.

Au cours des prochaines sections de ce chapitre, la méthodologie expérimentale complète est présentée. D'abord, les diverses étapes préliminaires seront décrites par une introduction à la taxinomie des coraux, aux méthodes d'acquisition des images et à la base de données. Ensuite, le système proposé sera décrit en détail avec une section couvrant la segmentation des spécimens, l'extraction de descripteurs, l'entraînement d'algorithmes intelligents pour la classification des spécimens et l'évaluation de la capacité de généralisation. Encore une fois, les résultats pertinents obtenus seront présentés suite à la description de la méthodologie. Ainsi deux problématiques majeures : la segmentation et la classification automatisées seront couvertes complètement par ce chapitre.

### **3.1 Introduction à la taxinomie des coraux**

La description taxinomique « est une branche des sciences qui vise à établir une classification systématique et hiérarchisée des taxons dans diverses catégories selon les caractères qu'ils ont en commun, des plus généraux aux plus particuliers ». Le taxon est quant à lui un « groupe d'êtres vivants qui ont des traits communs » (Office québécois de la langue française, 2010). La hiérarchie des taxons est définie par sept niveaux : règne, embranchement (*phylum*), classe, ordre, famille, genre, espèce. Tout comme la vie, cette science est en constante évolution, et encore plus depuis que les analyses ADN sont accessibles. De ce fait, on introduit parfois un huitième niveau, le sous-ordre, qui s'insère entre l'ordre et la famille.

Puisque nous cherchons toujours à servir la cause des Philippines, nous avons limité notre recherche aux coraux qui vivent dans ses eaux environnantes. Les coraux d'intérêt sont tous de l'ordre *scleractinia*, soit les coraux munis d'un squelette de calcite. En introduction, nous

mentionnions que les mers qui bordent les Philippines sont l'hôte de la plus grande diversité aquatique sur terre. Cette affirmation se confirme lorsque l'on considère qu'uniquement dans l'ordre *scleractinia*, nos collaborateurs ont répertorié 288 espèces de corail (Licuanan, 2008). Nos spécimens appartiennent donc tous aux mêmes règne, embranchement et classe. Voici une description de leur taxinomie et quelques faits divers et descriptions tirés de Wikipedia (Wikimedia Foundation, 2010) :

- Règne *animalia* : les animaux, êtres complexes pluricellulaires;
- Embranchement *cnidaria* : espèces animales aquatiques. Le terme *cnidaria* provient des animaux (cnidocytes ou cnidoblastes) qui possèdent des cellules urticantes. Par exemple, les méduses (*jellyfish*) sont une des deux formes qui appartiennent à cet embranchement. Les coraux (sous forme de polypes) sont la deuxième forme possible des cnidaires et sont parfois vénéneux. Il est intéressant de noter que la toxine non protéinaire la plus puissante répertoriée au monde est sécrétée par un corail d'Hawaii: le *palythoa toxica*, qui peut entraîner la mort en quelques minutes seulement (Rockefeller University, 2003);
- Classe *anthozoa* : comprends les coraux et les anémones de mer. Le nom *anthozoa* provient du Grec et signifie « animaux-fleurs ». Ce qui caractérise la symétrie inhérente à ces organismes;
- Ordre *scleractinia* : tel que mentionné précédemment, cet ordre concerne les coraux durs, munis d'un squelette. Ce sont les seuls types de coraux à former des récifs, d'immenses structures sous-marines de calcite qui prennent plusieurs milliers d'années à se former. Les récifs modifient la géométrie du benthos et sont à l'origine des lagons, ces étendues d'eau peu profondes bordant le littoral;
- Famille : plusieurs familles de corail sont présentes aux Philippines. Il y en a treize : *acroporidae*, *agariciidae*, *dendrophylliidae*, *euphyllidae*, *faviidae*, *fungiidae*, *merulinidae*, *mussidae*, *oculinidae*, *pectiniidae*, *pocilloporidae*, *poritidae* et *siderastreiiidae*;
- Genre et espèce: aux Philippines, on compte 288 espèces comprises dans 51 genres.

Cette vaste structure hiérarchique et complexe utilisée pour l'identification des coraux laisse croire que l'identification n'est pas une chose simple. En effet, la méthode manuelle reconnue scientifiquement exige un sens de l'observation très pointu et elle peut parfois être invasive, voir destructive pour le spécimen lorsqu'une collecte d'échantillon est nécessaire. Le guide électronique *Coral ID* de Veron et Stafford-Smith (2002) est un système expert qui permet, suite à une multitude de questions, d'identifier au niveau taxinomique le plus fin la nature d'un corail scléactinien parmi les 794 espèces disponibles. Cet outil est une version automatisée du protocole généralement reconnu pour identifier les espèces : la clé dichotomique. Le mécanisme de clé dichotomique est très répandu dans les sciences botanique et biologique. Il structure l'information quant aux différences notables entre les espèces vivantes ou fossilisées.

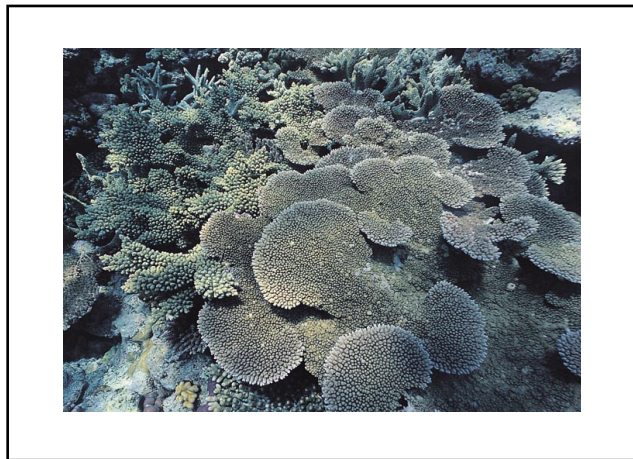


Figure 3.1 Un spécimen d'*acropora hyacinthus*.  
Adaptée de Veron et Stafford-Smith (2002)

Pour identifier un corail, prenons par exemple l'*acropora hyacinthus* (tel que présenté à la Figure 3.1) voici le cheminement possible tel que parcouru à l'aide de *Coral ID*.

Caractéristiques géographiques :

océan : Indo-pacifique

localisation indo-pacifique : mer rouge ou golf d'Aden

Caractéristiques morphologiques globales :

- colonial ou solitaire : colonial (multiples bouches ou corralites)
- attachement : attaché au substrat
- forme de croissance : ramifications
- croissance des ramifications : dominantes sur un plan bidimensionnel
- fusion des ramifications : dans un plan bidimensionnel seulement
- embout des ramifications : aplati
- ramifications primaires : absentes
- sous-ramifications : tridimensionnelles à disposition aléatoire

Caractéristiques du polype :

- cycles de septa : un seul cycle, ...

Caractéristiques du corallite (polype et son entourage) :

- taille de vallée : 1-5mm
- définition des corallites : centre distincts
- murs séparateurs : murs séparés
- isolation : individus séparés
- différentiation des corallites : présence de corallite axial
- squelette : complètement masqué

Caractéristiques temporelles :

- expansion de tissus : tentacules étendues le jour
- longueur des tentacules : inférieure à 10 mm
- projection d'autres tissus durant le jour : inférieure à 1 mm

Notez que pour chacune de ces questions, de multiples choix de réponses sont possibles. Seule une partie des caractéristiques relatives à l'*acropora hyacinthus* ont été présentées. Les caractéristiques du polype sont très nombreuses et nécessitent une étude approfondie pour bien saisir leur sens (elles utilisent une terminologie très spécialisée). Ce sont souvent ces

caractéristiques qui sont les plus discriminantes, malgré que leur matérialisation soit de taille inférieure au centimètre et parfois même inférieure au millimètre.

Finalement, même avec la méthode officiellement reconnue, il peut être difficile d'identifier précisément un spécimen. En fonction des caractéristiques choisies, il arrive souvent que de multiples espèces correspondent puisque pour une même espèce, les caractéristiques varient selon l'environnement immédiat du spécimen. Des photographies sont alors disponibles pour chaque espèce et permettent une comparaison visuelle pour finaliser l'identification. Plus encore, l'hybridation des coraux est possible et souvent rencontrée. Cette hybridation donne naissance à de nouvelles espèces qui ne sont pas documentées. Pour identifier les cas de figures complexes, *Coral ID* nous indique qu'un groupe de caractéristiques, appelé groupe tertiaire, existe. Ce groupe n'est pas encore disponible dans la version actuelle de cette clé dichotomique électronique.

Les caractéristiques tertiaires nécessitent davantage de connaissances sur la biologie des coraux et permettent d'identifier avec précision les cas les plus difficiles. Elles permettraient potentiellement de détecter l'hybridation ou les modifications rares attribuables au milieu de vie du spécimen. Puisque des notions approfondies en biologie marine sont nécessaires, nous n'avons pas poussé plus loin les recherches à ce sujet. Le but de cette section est simplement de montrer brièvement quelle est la méthode manuelle actuellement utilisée.

### **3.2 L'acquisition des images et la base de données**

Au cours de l'été 2009, nous avons obtenu l'autorisation d'utiliser les images fournies par la base de données de *Coral ID* (Veron et Stafford-Smith, 2002) pour effectuer nos recherches. Ces images ont été acquises, encore une fois en plongée sous-marine, par les Drs Veron et Stafford-Smith et certains collaborateurs. Cette fois, chaque image est bien identifiée par son espèce; ce qui sera un avantage pour nous, en termes de variabilité interclasse.



Toutefois, contrairement aux travaux du Dr Licuanan, aucun protocole scientifique n'a guidé la prise d'image. Les photos ont été acquises à des distances et à des angles de prise de vue variables, dans l'unique but de montrer la splendeur ou les caractéristiques inhérentes à un spécimen. Les photos montrent parfois une colonie de corail dans son ensemble, ou encore les détails des polypes. Cette variabilité dans les prises de vues complique les choses pour nous. De plus, chaque image présentée a été sélectionnée de façon à montrer toute la variabilité possible pour une espèce donnée. Dans un cadre encyclopédique, cette variabilité est symbole de qualité. Pour nous, la vocation encyclopédique de la banque d'image de *Coral ID* est un écueil qui ajoute une variabilité parmi le peu d'exemples qui représentent chaque espèce. En effet, les photographies de chaque espèce sont en faible nombre. Dans la portion de la base de données qui nous intéresse (les coraux des Philippines), on retrouve entre une et huit images; et plus généralement, quatre images sont utilisables par espèce (ce qui correspond à la médiane et au mode de la distribution des images des classes).

### **3.2.1 Extraction des images de *Coral ID***

Le logiciel *Coral ID* est une application Windows compatible à Microsoft Windows 98 ou Microsoft Windows XP. Lors de son exécution, le logiciel charge les *images JPEG* des coraux à partir du CD-ROM d'installation. Ce CD-ROM est donc la composante principale de la banque d'images utilisée par les prochaines expérimentations. Une deuxième composante est aussi nécessaire : la *base de données* qui permet d'attribuer à chaque image son espèce. Cette base de données est copiée sur le disque dur du système par l'environnement logiciel lors de l'installation ou de la première exécution. Ce fichier de base de données est de type Microsoft Access et maintient l'association du genre et de l'espèce d'un spécimen au code alphanumérique utilisé en guise de nom de fichier image. Notez que pour rendre plus conviviale l'accès à l'espèce présente sur chaque image, le contenu de la base de données Access a été extrait et consigné dans un fichier CSV (*comma separated values*, valeurs séparées par des virgules). Ce format de fichier est facile d'utilisation dans plusieurs outils tels MATLAB® et OOo (OpenOffice.org).

Voici quelques détails techniques ainsi que les chemins d'accès aux *images JPEG* et à la *base de données Access* :

- *Images JPEG (Joint photographic experts group)* : CD-ROM/Lucid/Coralid/images/  
Les images sur le CD-ROM sont structurées dans une dizaine de sous arborescences relatives au répertoire .../images/, numérotées comme suis : 001-100/Large/, 101-200/Large/, ..., 901-1000/Large/. Les images sont identifiées par une série de numéros également.
- *Base de données Access* : RACINE/Program Files/Lucid/Cache/coralid/Coralid.key  
Le fichier « key » consigne les données permettant d'associer les chemins de fichiers des images à un nom d'espèce particulier. Pour y accéder, le fichier « Coralid.key » peut être renommé « Coralid.mdb ». Une fois ouvert dans Microsoft Access, ce fichier de base de données présente une dizaine de tables. La table « Taxpictures » contient l'information importante : l'espèce est notée dans le champ « Taxname » et le chemin d'accès de l'image (relatif au répertoire .../images/ du CD-ROM) est consigné sous l'intitulé « Picture ».

### 3.2.2 Sélection des images de *Coral ID*

Certaines images de l'environnement *Coral ID* ne sont pas pertinentes à l'entraînement d'un système automatisé de reconnaissance corallienne. Avant toute chose, ces images ont été éliminées. Pour y arriver, nous avons isolé trois groupes d'images impertinentes à l'aide des suffixes alphabétiques juxtaposés aux noms des fichiers. Le suffixe « BW » correspond à des images en noir et blanc du squelette de calcite d'un spécimen, le suffixe CD correspond à des cartes géographiques relatives à l'espèce et enfin, le suffixe « D » correspond à des dessins à main levée des spécimens. Ces images ne correspondent pas au problème de reconnaissance des coraux dans leur milieu naturel et ne seront pas considérées. L'implémentation à cette étape a été effectuée à l'aide de MATLAB® et du fichier CSV de base de données.

De plus, tel que mentionné en 3.1, nous limiterons la recherche aux coraux des Philippines. Notre collaborateur de l'Université De La Salle nous a fourni une liste exhaustive des 288 espèces répertoriées sous les sites d'étude (Licuanan, 2008). Pratiquement toutes ces espèces (286 exactement), se retrouvent dans l'encyclopédie de Veron (2000) et complémentaiement dans la clé dichotomique *Coral ID* de Veron et Stafford-Smith (2002). Nous disposons donc de quelques images de ces espèces d'intérêt (voir la Figure 3.2 pour évaluer les quantités). L'implémentation à cette étape a été conduite à l'aide d'un outil de comparaison de fichiers textes : FileMerge sous la suite de développement d'Apple. Cette comparaison des listes d'espèces était nécessaire pour repérer les coquilles, puisque la liste des espèces des Philippines a été saisie manuellement par notre collaborateur. Enfin, pour construire la liste exhaustive des chemins de fichiers images et des espèces correspondantes, les données ont été importées dans un moteur de base de données, OpenOffice.org Base. Une requête SQL a été utilisée pour extraire les images pertinentes. Suite à l'élimination des données non considérées, la base de données CSV compte 1079 images, réparties dans 286 espèces.

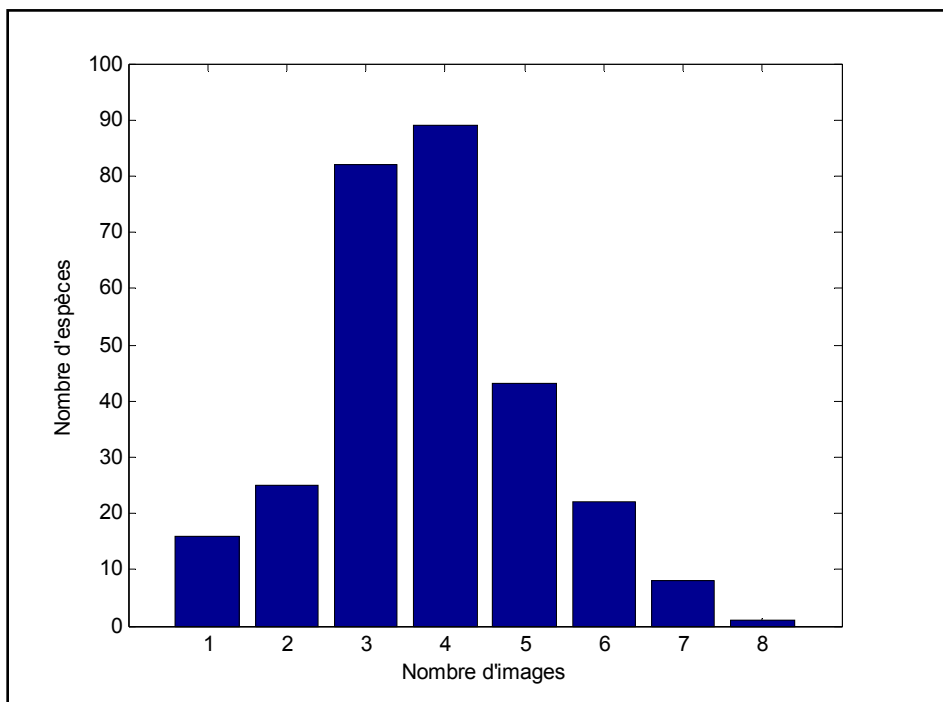


Figure 3.2 Répartition des espèces en fonction de la quantité d'images de *Coral ID*.

### 3.3 L'échantillonnage des données et les méthodes de validation

L'échantillonnage et la validation sont deux points critiques de l'élaboration d'une expérience. L'échantillonnage permet de concevoir les expérimentations qui valideront nos hypothèses et de préciser la quantité de données considérée. Enfin, les méthodes de validation s'assurent de bien diviser les données pour accomplir indépendamment la validation des métaparamètres, l'apprentissage des classificateurs et le test en généralisation. De plus, le processus de validation se doit de diviser correctement la base de données en segments représentatifs et indépendants pour assurer un nombre suffisant de répliques comparables statistiquement.

#### 3.3.1 L'échantillonnage des données

En plus de la sélection préalable des classes à l'étude décrite en 3.2.2, divers sous échantillonnages seront testés. Le premier but de cet échantillonnage secondaire est de sélectionner les données pertinentes de façon à répondre aux questions soulevées par nos hypothèses de travail.

Tout d'abord, il est important de constater qu'avec un nombre aussi faible d'images par classes, il sera difficile de résoudre la problématique complète, qui comprend 286 espèces. Pour limiter les temps de préparation des données (segmentation manuelle, etc.), pour limiter les temps de calcul et pour éviter d'effectuer les tests sur une problématique trop vaste pour le nombre d'échantillons disponibles; nous avons limité notre recherche aux 31 classes les plus représentées de la base de données. C'est à dire, les 31 espèces de corail des Philippines (10,8 % des espèces disponibles) qui possèdent entre 6 et 8 images (soit la partie de droite de l'histogramme de la Figure 3.2). À première vue, les espèces choisies représentent bien la situation : on remarque immédiatement que le genre *acropora* est fortement représenté, comme observé aux Philippines. Le Tableau 3.1 présente cet échantillonnage et la répartition des 196 images issues de *Coral ID* concernées (soit 18,2 % de la quantité d'images totale).

Cette fois, l'implémentation de la sélection a été effectuée à l'aide d'un tableur : OpenOffice.org Calc.

Tableau 3.1 Les espèces considérées par l'échantillonnage

No	Famille	Genre	Espèce	Qté. d'images
0	acroporidae	acropora	cerealis	7
1	acroporidae	acropora	cytherea	6
2	acroporidae	acropora	florida	6
3	acroporidae	acropora	glauca	6
4	acroporidae	acropora	hyacinthus	6
5	acroporidae	acropora	nobilis	6
6	acroporidae	acropora	palifera	7
7	acroporidae	acropora	subulata	6
8	acroporidae	astreopora	myriophthalma	6
9	acroporidae	montipora	aequituberculata	8
10	acroporidae	montipora	tuberculosa	7
11	dendrophylliidae	turbinaria	stellulata	6
12	euphyllidae	euphyllia	ancora	6
13	euphyllidae	physogyra	lichtensteini	6
14	euphyllidae	plerogyra	sinuosa	6
15	faviidae	cyphastrea	serailia	6
16	faviidae	echinopora	lamellosa	6
17	faviidae	favites	complanata	6
18	faviidae	goniastrea	australensis	7
19	faviidae	montastrea	annularis	7
20	faviidae	montastrea	curta	6
21	faviidae	platygyra	lamellina	6
22	fungiidae	cycloseris	cyclolites	6
23	mussidae	lobophyllia	hemprichii	6
24	oculinidae	echinophyllia	aspera	7
25	oculinidae	galaxea	astreata	6
26	pocilloporidae	pocillopora	damicornis	6
27	pocilloporidae	stylophora	pistillata	7
28	poritidae	goniopora	lobata	6
29	poritidae	porites	cylindrica	7
30	poritidae	porites	lobata	6

L'échantillonnage des 31 classes les plus représentées des coraux philippins sera ensuite sous échantillonné pour résoudre les deux problèmes envisagés : la segmentation et la classification.

### **3.3.1.1 Évaluation de la performance en segmentation**

Premièrement, l'échantillon complet sera utilisé pour entraîner et valider un algorithme de reconnaissance de formes appliqué à la segmentation automatisée du corail selon la procédure de validation croisée explicitée en 1.3.2.3. Pour ce qui est de la segmentation, ce sera le seul test effectué. Notez que nous considérerons la problématique de segmentation et de classification indépendamment.

### **3.3.1.2 Évaluation de la capacité de classification considérant deux métaclases**

Ensuite, pour résoudre le problème de classification, le premier test envisagé regroupe les espèces en deux catégories : *acropora* et *non-acropora*. Ce test a pour but de valider l'approche avant de procéder à l'évaluation de la quantité de données nécessaires à résoudre toute la problématique. Comme mentionné au CHAPITRE 2, une distinction est souvent faite par les chercheurs en biologie entre les spécimens *acropora* et *non-acropora* : les coraux du genre *acropora* sont très nombreux et selon Licuanan (2001), on distingue un corail de ce genre par son évidente structure coralliaire (*corallite*) axiale au sommet entourée d'autres structures radiales. Ainsi, nous pourrions voir les performances atteintes lorsque chaque classe est renseignée par un lot important d'échantillons.

### **3.3.1.3 Évaluation de l'effet du nombre de classes sur les performances**

Par la suite, l'échantillonnage des 31 espèces de coraux philippins sera utilisé pour valider la quantité de classes qu'il est possible de départager correctement considérant toutes les données disponibles. Au cours de ce test, huit réplifications sont exécutées en variant aléatoirement le groupe de classes considéré (de 2 à 6 classes). Ainsi, la difficulté à

départager les classes les unes des autres (considérant un groupe prédéterminé) sera évaluée de huit façons différentes. Avec ces répliques, l'effet inhérent aux variabilités intraclasse et interclasses particulières aux espèces choisies sera amoindri. La sélection aléatoire des classes a été fixée à l'avance et est présentée par le Tableau 3.2. Dans ce tableau, le numéro identifiant chaque classe est conséquent avec celui présenté au Tableau 3.1.

Tableau 3.2 Groupes de classes échantillonnés

Nb. de classes	Groupes
2	{12, 20}, {6, 19}, {14, 13}, {5, 26}, {21, 25}, {16, 3}, {29, 15}, {2, 7}
3	{12, 20, 6}, {19, 14, 13}, {5, 26, 21}, {25, 16, 3}, {29, 15, 2}, {7, 9, 11}, {23, 27, 17}, {24, 22, 0}
4	{12, 20, 6, 19}, {14, 13, 5, 26}, {21, 25, 16, 3}, {29, 15, 2, 7}, {9, 11, 23, 27}, {17, 24, 22, 0}, {28, 4, 18, 30}, {8, 1, 10, 12}
5	{12, 20, 6, 19, 14}, {13, 5, 26, 21, 25}, {16, 3, 29, 15, 2}, {7, 9, 11, 23, 27}, {17, 24, 22, 0, 28}, {4, 18, 30, 8, 1}, {10, 12, 20, 6, 19}, {14, 13, 5, 26, 21}
6	{12, 20, 6, 19, 14, 13}, {5, 26, 21, 25, 16, 3}, {29, 15, 2, 7, 9, 11}, {23, 27, 17, 24, 22, 0}, {28, 4, 18, 30, 8, 1}, {10, 12, 20, 6, 19, 14}, {13, 5, 26, 21, 25, 16}, {3, 29, 15, 2, 7, 9}

### 3.3.1.4 Évaluation de l'effet de la quantité de données par classe sur les performances

Enfin, la dernière expérience présentée montrera comment le système se comporte lorsqu'on fait varier la quantité de données d'apprentissage, considérant un nombre de classes à départager fixe. Huit tests ont été conduits, cette fois en variant à chaque itération la quantité de données utilisée pour entraîner les classificateurs. La quantité de données a été mesurée en

terme de surface (l'aire en pixels, la quantité totale de pixels utilisée) de la zone image considérée par le classificateur pour chaque classe. La performance, considérant une surface image par classe de  $1 \times 10^5$  à  $8 \times 10^5$  pixels, a été évaluée (où  $8 \times 10^5$  correspond approximativement à l'aire maximale admise par la classe qui est la moins bien représentée).

Pour chacune des itérations, huit réplifications ont été conduites en variant aléatoirement les segments d'images utilisés pour constituer la surface désirée. Ainsi, l'effet de la variabilité inter et intraclasse associée à chaque segment a été minimisé. Ce processus (8 surfaces  $\times$  8 réplifications) a été répété pour chaque groupe de 2 à 6 classes à départager. Les classes considérées pour effectuer ce test sont issues du test précédent. Par exemple, le groupe de deux classes choisi pour exécuter la procédure est le groupe qui a produit la performance médiane au test précédent (en 3.3.1.3). Ainsi, les performances obtenues seront plus représentatives de la réalité, sans toutefois nécessiter l'exécution de ce test sur chaque groupe de classes admissible.

### **3.3.2 La méthode de validation**

Pour réaliser les diverses expériences, la méthode de validation croisée sera utilisée. Cette méthode permet d'effectuer un nombre satisfaisant de réplifications quasi indépendantes, tout en prenant compte du faible nombre d'échantillons disponibles. Avec un nombre aussi faible que huit images par espèce, la réplification simple, qui implique une division préalable des données et met de côté un nombre de sous-ensembles égal au nombre de réplifications, n'est pas une solution envisageable. Pour limiter l'impact de ce problème, la validation croisée a été implémentée sur de multiples segments d'images, extraits de chacune des 6 à 8 images complètes disponibles par espèce. Cette procédure de présegmentation est présentée en 3.4.2.1.

Notre procédé de validation croisée utilise huit partitions qui sont stratifiées selon les classes. La stratification assure une répartition similaire des classes dans chaque bloc de validation. Ainsi, la probabilité a priori d'obtenir l'une ou l'autre des classes lors d'une pige aléatoire



dans n'importe quel bloc reste similaire à celle observée pour une pige parmi l'ensemble des données. Cette propriété statistique est importante, puisque la plupart des paradigmes de reconnaissance de formes (RDF) utilisent la probabilité à priori pour guider le processus d'apprentissage. Les huit réplifications seront donc homogènes à cet égard.

Nous avons choisi huit blocs pour effectuer notre validation croisée pour plusieurs raisons. En premier lieu, la faible quantité de données disponibles nous oblige à maintenir un petit nombre de réplifications pour assurer la stabilité des tests. Il est impératif qu'assez de données soient disponibles pour entraîner l'algorithme intelligent, pour qu'une grande partie de la variabilité inhérente aux classes soit captée et prise en compte par les frontières de décision. Enfin, la partition réservée aux tests doit également être suffisamment grande pour que tous les cas de figure importants soient mis à l'essai. Sans quoi, chaque réplification ne pourrait être comparée adéquatement : la variance des performances observées sera trop importante pour conclure sur la qualité d'un système. Une autre raison nous pousse à choisir huit blocs pour valider nos systèmes est la parallélisation des calculs. Les ordinateurs utilisés pour exécuter les expériences possèdent huit cœurs, le nombre de réplifications choisi (qui est multiple de huit) permet de maximiser la performance en exécution parallèle en utilisant simultanément les huit cœurs disponibles à plein rendement. L'heuristique habituelle, la validation croisée à dix itérations, est moins efficace dans ce contexte puisqu'après les huit premières itérations seuls deux processeurs sont utilisés pour compléter l'expérience. Enfin, selon nos tests préliminaires, la variance observée au cours des huit réplifications est assez faible, ce qui résulte en un intervalle de confiance satisfaisant malgré le faible nombre d'expériences. Inutile alors d'accroître ce nombre à 16 ou 24 réplifications.

Enfin, les huit blocs sont utilisés comme suis : un bloc est d'abord utilisé pour valider les métaparamètres à utiliser par l'algorithme de reconnaissance de forme, les six blocs suivants sont utilisés pour entraîner l'algorithme intelligent et enfin, le dernier bloc est utilisé à titre de test, pour extraire les performances du système étudié. À chacune des huit itérations, les blocs utilisés pour la validation, l'entraînement et le test sont interchangés selon la procédure de rotation habituelle (voir la section 1.3.2.3 pour plus de détails sur la validation croisée).

Pour finalement évaluer et comparer les performances d'un système, la moyenne est l'estimateur statistique que nous utiliserons. Pour bien situer la performance, un intervalle de confiance sera calculé sur chaque moyenne à partir de la distribution de Student à sept degrés de liberté. Les intervalles donneront un aperçu visuel (suite à une représentation graphique) de la performance et de la stabilité de divers systèmes. Pour plus d'information sur le calcul de ces intervalles de confiance, voir la section 1.3.6.4.

*Note : Puisque le nombre d'échantillons disponibles est très limité, nous considérerons la problématique de la segmentation et de la classification indépendamment. En d'autres termes, les mêmes données seront utilisées pour entraîner le système de segmentation supervisée et le système de classification supervisée. Dans un contexte de développement pour la mise en service du système, l'entraînement de ces deux entités devrait être effectué sur deux segments indépendants de la base de données pour éviter le surapprentissage. Le tout, considérant que le but ultime est de relier ces deux systèmes en cascade.*

### **3.4 La segmentation des spécimens**

Les photographies des spécimens de *Coral ID* ne sont pas segmentées à l'origine. En d'autres termes, chaque image présente en son centre un sujet principal (le corail d'une espèce en particulier) et en périphérie, un arrière-plan composé de divers objets (d'autres coraux, des algues, du sable ou de l'eau bleutée à l'horizon). Avant tout traitement, la zone où se trouve le sujet d'intérêt de chaque image doit être identifiée; c'est ce processus que nous appelons la *segmentation*. Son automatisation est une des deux principales contributions de ce travail.

Ce traitement préalable des données est important : la première étape pour arriver à identifier la nature d'un objet est sans contredit la localisation de celui-ci dans l'espace de description. La segmentation permet d'établir la zone de chaque image qui doit être étudiée lors de l'extraction de caractéristiques pour la classification. Ainsi, les données extraites pour l'identification de l'espèce d'un corail ne seront pas biaisées par son environnement immédiat. Le but ultime est d'accomplir cette tâche automatiquement par un mécanisme de

vision artificielle (VA). Toutefois, pour permettre l'entraînement et l'évaluation d'un tel algorithme de segmentation intelligent et complètement automatisé un lot de données segmentées manuellement est initialement nécessaire.

### **3.4.1 La segmentation manuelle**

Pour établir une segmentation de référence des images de *Coral ID* un outil de segmentation adapté à cette problématique a été conçu de toutes parts. Cet outil permet à son utilisateur de parcourir les images échantillonnées en 3.3 et de superviser une procédure de segmentation semi-automatisée. Selon les conditions présentes dans une image, l'outil n'arrive pas toujours à segmenter avec précision tout le corail au premier essai. Des corrections manuelles peuvent alors être apportées à cette segmentation initiale. En combinant cet outil logiciel à une tablette graphique avec stylet, plusieurs centaines d'images de corail peuvent être segmentées avec précision en quelques heures seulement. Les prochaines pages présenteront l'outil développé et ses fonctionnalités.

Puisque plusieurs outils de traitement d'images sont déjà implémentés dans MATLAB®, l'outil de segmentation manuel et son interface ont été développés sous cet environnement. L'interface graphique donne accès à plusieurs contrôles, outils et modes de fusion. Tous ces items sont accessibles par la barre d'outils et également par des raccourcis clavier. L'outil de segmentation tient pour acquise la présence du fichier CSV et du répertoire d'images du CD-ROM de *Coral ID*. Suite au chargement, l'interface graphique est affichée à l'écran ainsi que la première image de la séquence à segmenter tel qu'illustré à la Figure 3.3.

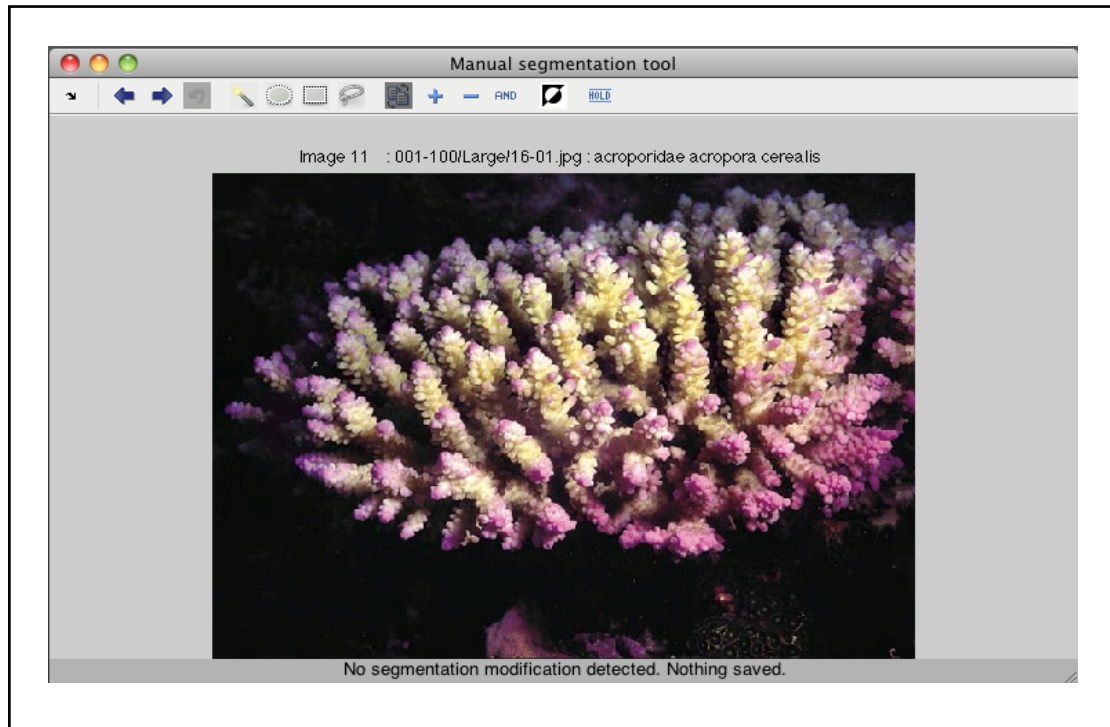


Figure 3.3 Outil de segmentation manuel (état initial).  
L'image de corail a été adaptée de *Coral ID* (Veron et Stafford-Smith, 2002)

Déjà à cette étape, une opération cruciale a été effectuée : la segmentation automatique du pourtour de l'image. Dans l'environnement graphique de *Coral ID*, chaque image est présentée sur fond gris et pour uniformiser la taille des images les auteurs ont ajouté un arrière-plan gris et une bordure d'image noire (voir Figure 3.4). Ces segments d'images inutilisables sont éliminés automatiquement suite au chargement. Ce prétraitement est effectué à l'aide de méthodes d'analyse d'images simples : l'indice RVB (rouge-vert-bleu) du pourtour gris à éliminer est constant pour toutes les images, il est donc évident de le segmenter.

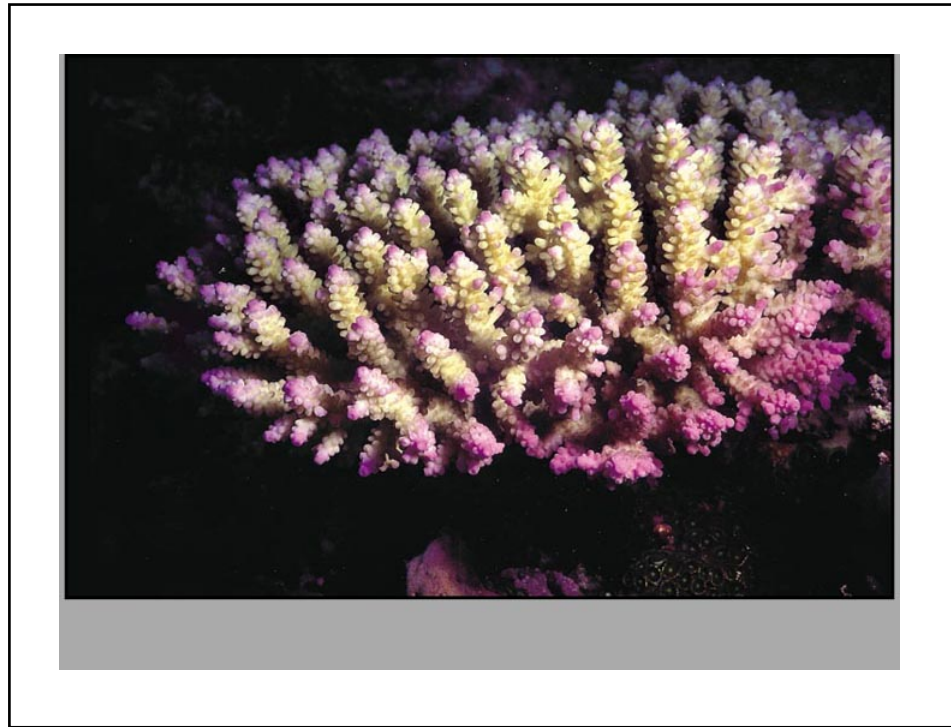


Figure 3.4 Image originale de *Coral ID*.  
L'image de corail a été tirée de *Coral ID* (Veron et Stafford-Smith, 2002)

Suite à la segmentation du contour (qui est effectuée de façon transparente à l'utilisateur), l'utilisateur peut démarrer la procédure de segmentation semi-automatisée du spécimen. Pour y arriver, plusieurs outils sont disponibles. La Figure 3.5 illustre les divers outils, modes et contrôles de l'interface et leur raccourci clavier respectif sont présentés entre parenthèses. Les prochaines pages en détailleront le fonctionnement.

Les premiers contrôles permettent de naviguer dans la banque d'images. Chaque fois que l'utilisateur se déplace dans la liste des images à segmenter, le résultat de la segmentation actuelle est sauvegardé automatiquement. Pour segmenter une image, l'utilisateur dispose d'un choix de quatre outils communs à tous les logiciels d'édition d'images. La baguette magique (*magic wand*) est l'outil de choix pour commencer la segmentation. Cet outil permet de sélectionner un point de référence sur le corail et suite à la sélection, tous les pixels similaires en couleur et en intensité lumineuse sont automatiquement sélectionnés.

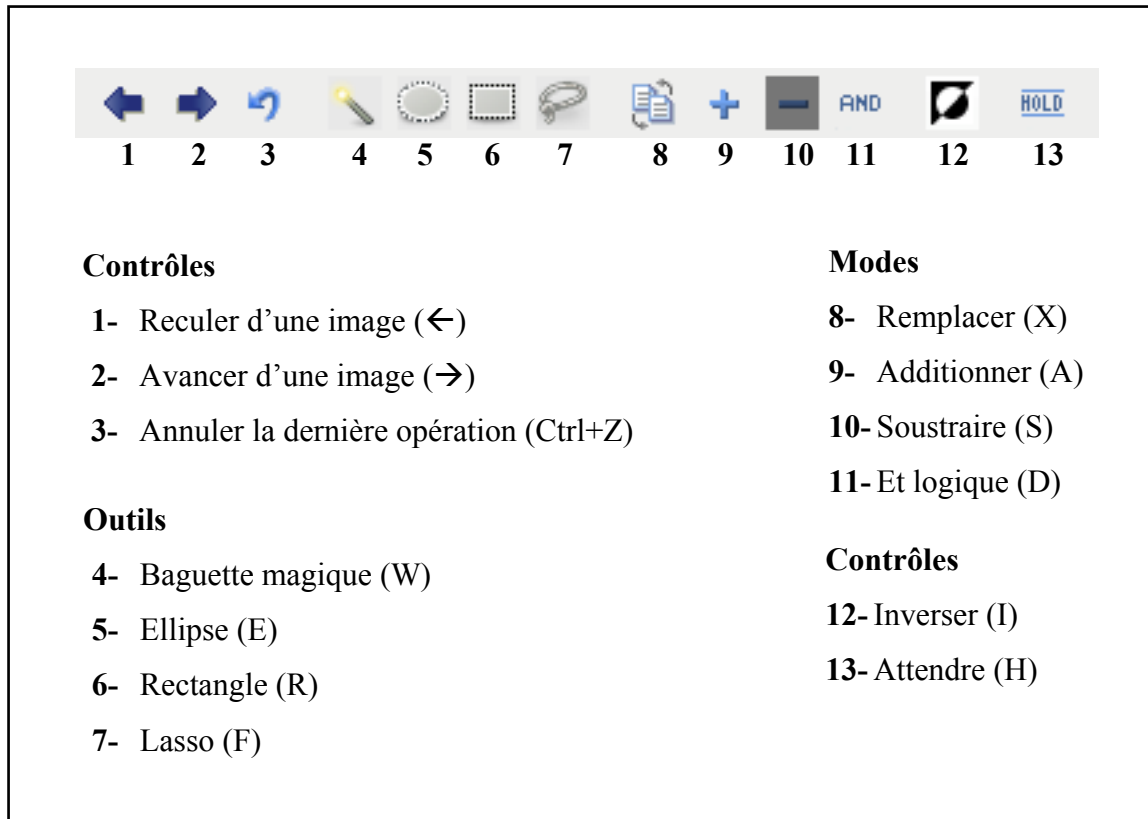


Figure 3.5 Les outils, modes et contrôles de l'outil de segmentation.

Avant de sélectionner un point sur le corail, certains paramètres peuvent être ajustés, comme l'illustre la Figure 3.6. Le premier paramètre : la tolérance (*wand sens*) permet d'ajuster la quantité de pixels images qui répondront au critère de similarité. Plus cette tolérance est grande, plus nombreux seront les pixels sélectionnés, lors de la segmentation d'une image naturelle; puisqu'une plus grande différence entre ces pixels et la référence sera tolérée. Le rayon (*radius*) permet de déterminer la quantité de pixels de référence qui seront choisis autour du point cliqué. Plus le rayon est grand, plus l'échantillon de référence sera important, ce qui accroît généralement la quantité de pixels qui seront sélectionnés également. Enfin, le mode permet de choisir si la sélection s'étend sur toute l'image (*global*), ou sur l'unique zone jugée similaire et contigüe avec la référence choisie (*local*). Notez qu'en tout temps, il est possible d'appuyer sur la touche d'échappement pour désactiver un outil.

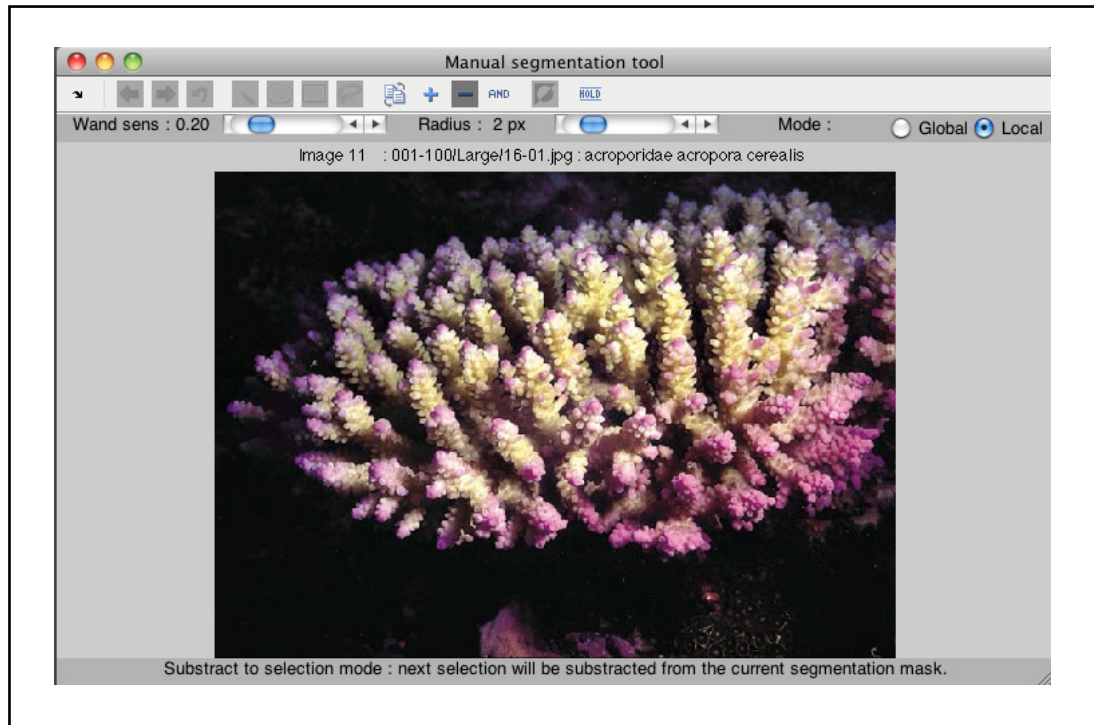


Figure 3.6 Outil de segmentation (configuration de la baguette magique).  
L'image de corail a été adaptée de *Coral ID* (Veron et Stafford-Smith, 2002)

Dans la plupart des cas, l'outil « baguette magique » effectue une segmentation acceptable en un seul clic. La Figure 3.7 illustre le résultat de la segmentation initiale. L'affichage de la segmentation indique par une mince ligne rouge le contour de la zone segmentée et ombrage la zone qui doit être considérée comme n'appartenant pas au sujet. Notez toutefois qu'une astuce importante a été utilisée pour segmenter ce corail. L'arrière-plan de la scène est plus uniforme que le corail lui-même. Dans ce cas, il est avantageux d'utiliser le mode « soustraire à la sélection » et d'utiliser l'outil « baguette magique » en sélectionnant un pixel de l'arrière-plan à titre de référence. Ainsi, la segmentation est beaucoup plus précise et une seule opération permet d'effectuer une grande partie du travail. Évidemment cela est vrai pour une image dont le fond est assez constant, mais moins vrai pour une image qui comporterait d'autres coraux en arrière plan.

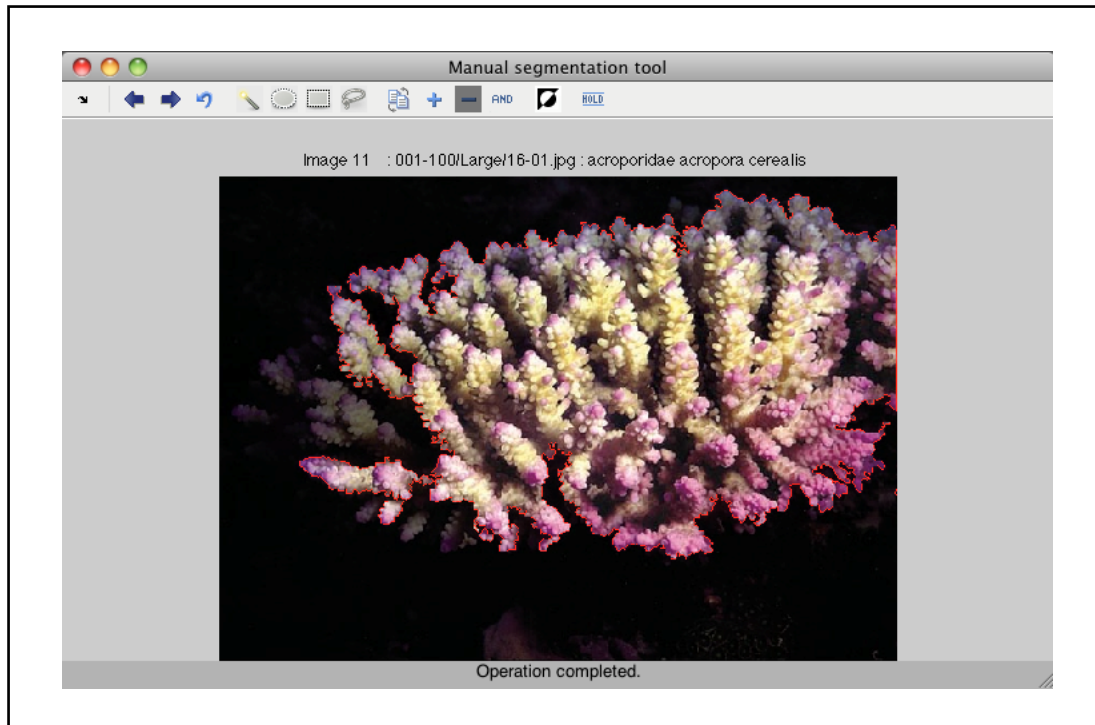


Figure 3.7 Outil de segmentation (après la sélection).  
L'image de corail a été adaptée de *Coral ID* (Veron et Stafford-Smith, 2002)

Le résultat initial est satisfaisant. Toutefois, une petite zone du corail (à gauche) n'a pas été prise en compte. Pour bien illustrer le fonctionnement de l'outil, nous allons continuer cette segmentation jusqu'à l'obtention d'une sélection parfaite. Pour ce faire, une deuxième utilisation de l'outil « baguette magique » permet de corriger le petit défaut. Notez qu'il faut alors sélectionner le mode « additionner à la sélection » pour faire en sorte que le prochain clic, qui sera fait sur une partie du corail cette fois, ajoute le morceau désiré à la sélection. Suite à quelques exécutions de la commande « baguette magique », nous obtenons le résultat de la Figure 3.8. Remarquez qu'en sélectionnant d'autres pixels du corail, un artefact indésirable est apparu dans la zone supérieure de l'image. Dans ce cas, il est facile de corriger la situation en utilisant un autre outil, tels le rectangle, l'ellipse ou encore le lasso de sélection.



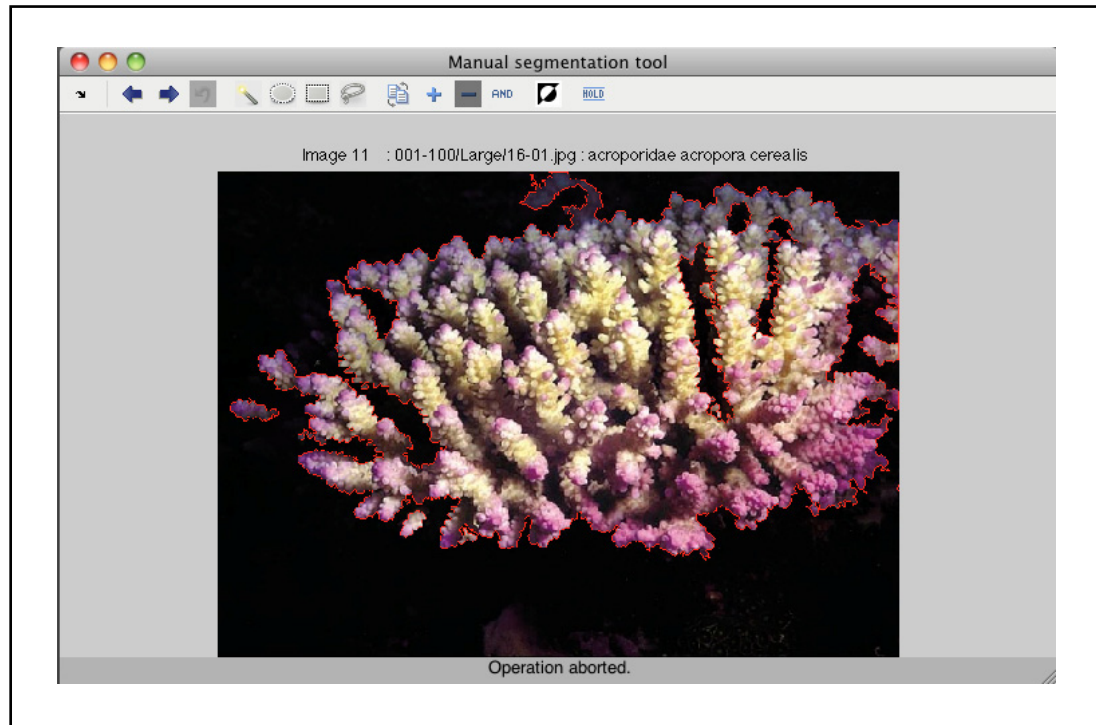


Figure 3.8 Outil de segmentation (affinement de la sélection).  
L'image de corail a été adaptée de *Coral ID* (Veron et Stafford-Smith, 2002)

La Figure 3.9 illustre le fonctionnement de l'outil « lasso », pendant son tracé. Le lasso permet de tracer avec la souris, ou plus préférablement avec une tablette graphique, une zone de forme personnalisée qui peut être, selon le mode choisi, ajoutée ou soustraite à la sélection. Cet outil flexible permet de segmenter les coraux qui se trouvent dans un environnement très bruité ou très similaire au spécimen lui-même. Un tel scénario est rare. Toutefois, ces cas de figure rendent très difficile l'utilisation de la baguette magique qui aurait tendance à sélectionner tous les pixels de l'image. C'est pourquoi d'autres outils, tel le lasso, sont disponibles.

La sélection du mode « soustraire à la sélection » et la saisie d'un tracé fermé avec l'outil « lasso » permettent d'éliminer l'artéfact indésirable. Dès que la pression sur le bouton de la souris ou sur le stylet est relâchée, l'outil « lasso » ferme le tracé en reliant le point de départ et le point de fin par une ligne droite et effectue l'opération de segmentation demandée. La Figure 3.10, illustre le résultat final obtenu.

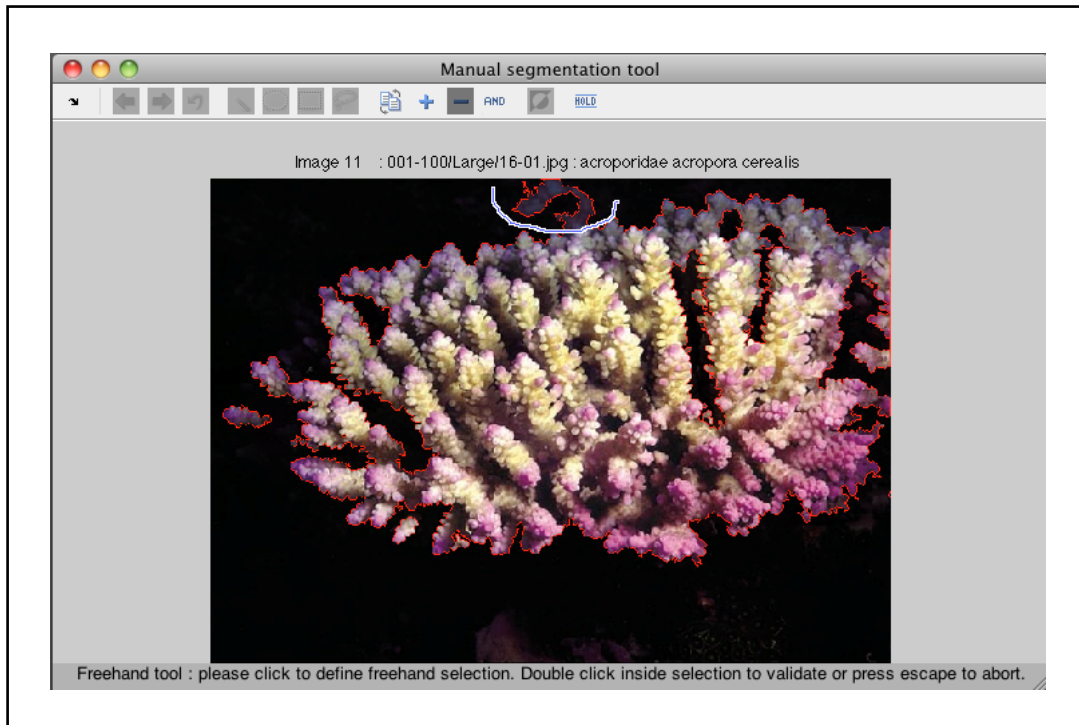


Figure 3.9 Outil de segmentation (utilisation du lasso).  
L'image de corail a été adaptée de *Coral ID* (Veron et Stafford-Smith, 2002)

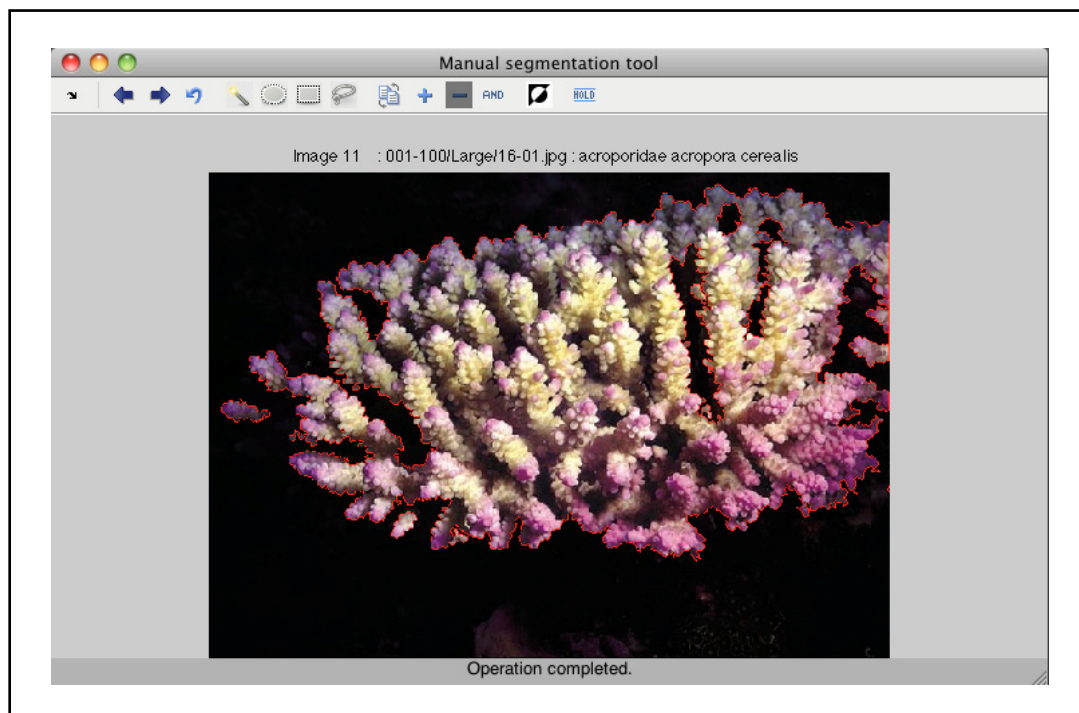


Figure 3.10 Outil de segmentation (résultat final).  
L'image de corail a été adaptée de *Coral ID* (Veron et Stafford-Smith, 2002)

Au cours de l'exemple précédent, certains outils, modes et contrôles n'ont pas été présentés. Au niveau des contrôles, notez qu'il est possible d'annuler la dernière opération de segmentation effectuée. Le bouton d'annulation est pratique lors de l'utilisation de la baguette magique, puisque le résultat obtenu peut être difficile à prévoir de prime abord. Il existe également une commande pour inverser la sélection en cours. Ce bouton permet en un seul clic d'invertir la zone en surbrillance avec celle qui est ombragée. Finalement, la commande « attendre » (*hold*) est une option qui permet de personnaliser le comportement de l'interface suite à la saisie d'un point de référence (baguette magique), d'un rectangle, d'une ellipse et d'un tracé au lasso. Lorsque la commande « attendre » est activée, l'outil de segmentation attendra une confirmation de l'utilisateur avant de lancer la segmentation. Ainsi il est possible d'effectuer de fins ajustements sur la position et la taille des zones sélectionnées par les divers outils suite à une saisie initiale approximative. Pour confirmer la position et la taille d'une sélection, il suffit alors de double-cliquer dans la sélection à l'écran.

Enfin, lorsque la segmentation à l'écran est adéquate, il suffit de passer à l'image suivante ou encore de quitter l'outil de segmentation pour que le résultat soit sauvegardé automatiquement. La sauvegarde s'effectue sous la forme d'un masque binaire matriciel. Tous les nombres « 1 » de la matrice correspondent à la zone segmentée et tous les « 0 » correspondent à l'arrière-plan qui doit être éliminé. Cette matrice binaire est compressée sans perte dans un fichier au format PNG (*portable network graphics*). Par la suite, chaque fois que l'utilisateur charge une image avec l'outil de segmentation, ce masque binaire est également chargé, s'il est disponible. La séance de segmentation peut alors être reprise où elle avait été laissée précédemment.

Pour conclure cette section, voici une description mathématique plus détaillée de l'algorithme de la baguette magique. L'outil « baguette magique » utilise simplement la distance euclidienne entre l'intensité  $i$  des divers pixels de l'image et des  $n$  pixels de référence d'intensité  $r$ . Pour définir le masque de segmentation binaire  $m$ , la distance totale sur toutes les dimensions couleur  $d$  doit être inférieure à la tolérance  $t$ . L'équation (3.1) résume cette relation sous forme d'opérations matricielles. La distance est calculée dans

l'espace de couleur RVB, tout simplement. Diverses représentations de la couleur et de la luminosité, donc HSV (*hue saturation value*) et Lab (luminance, a, b), ont été testées durant la phase de développement de l'outil. Nous avons toutefois constaté que l'espace RVB est un choix simple, efficace et est le plus couramment utilisé par les logiciels de traitement de l'image. La tolérance  $t$  s'exprime donc simplement comme la norme d'un vecteur dans l'espace normalisé de l'intensité colorimétrique RVB.

$$dist_n = \sum_d (i_d - r_{n,d})^2 ; m_n = dist_n \leq t^2 ; \boxed{m = m_1 || m_2 || \dots || m_n} \quad (3.1)$$

### 3.4.2 La segmentation automatique supervisée

Maintenant qu'un lot d'images segmentées est disponible, il sera possible d'entraîner un algorithme d'intelligence artificielle (IA) pour accomplir le travail automatiquement. Cette section traitera des diverses méthodes qui sont utilisées pour s'acquitter de cette tâche de RDF. Il est toutefois important de considérer le contexte de cette segmentation automatisée. Le but ultime de ce travail est de fournir un outil qui est capable de - 1 - repérer les coraux dans une image numérique bidimensionnelle (par segmentation/détection) et - 2 - d'identifier au niveau taxinomique chaque spécimen de corail repéré dans ces images (par classification). En revanche, la banque d'images (*Coral ID*) utilisée pour faire la preuve de concept n'a pas été acquise de façon à faciliter l'atteinte de ces objectifs ambitieux. Dans chaque image de *Coral ID* un sujet principal occupe près des trois quarts du champ image, est bien au foyer et est identifié au niveau taxinomique. Notre segmentation manuelle isole seulement ce sujet photographique. L'arrière-plan qui est éliminé au cours du traitement manuel de 3.4.1 est parfois tapissé d'autres coraux, dont l'identification n'est pas disponible.

Ces coraux en arrière-plan ajoutent à la confusion que doit gérer l'algorithme intelligent de détection du corail. Malgré tout, nous avons choisi d'utiliser l'information issue de la segmentation manuelle du sujet photographique de chaque image pour entraîner le système. La raison de ce choix s'exprime simplement : la répartition des classes « corail » et « non-

corail » est très biaisée par la vocation encyclopédique de notre base de données. En termes plus simples, Dr Veron n'a sélectionné que les images qui mettent en évidence une espèce de corail particulière pour créer *Coral ID*. Ainsi, le corail étudié est généralement centré, couvre la majorité de la zone image et chaque image admet un arrière-plan de nature inconnue en périphérie du sujet. Dans ces conditions, nous n'aurions eu que très peu d'information sur la classe « non-corail », si la segmentation avait été accomplie dans le but d'éliminer véritablement la portion de chaque image qui ne peut être considéré comme un corail. Ce choix est d'autant plus pratique considérant qu'il permet d'utiliser la même segmentation manuelle des images pour entraîner le détecteur de corail et le classificateur de corail.

Enfin, malgré le défi que représente la détection du corail (ou plutôt du sujet photographique) dans ce contexte, notre hypothèse de travail reste plausible : étant donné l'information véhiculée par le point focal de l'image et l'intention du photographe (sur chaque image, il nous apparaît très évident de départager le sujet de l'arrière-plan). De plus, considérant également que ce contexte complique le problème de segmentation, nous posons également l'hypothèse que les conclusions que nous pourrions tirer de cette analyse sont transférables à celle d'un « véritable » système de détection de corail. Lorsqu'une base de données adéquate sera disponible, le problème de segmentation automatisée (ou de détection, selon la terminologie utilisée) sera sans doute plus facile à résoudre puisqu'aucune confusion entre la nature des objets à classifier ne sera présente dans ces données (aucun corail ne se retrouvera étiqueté comme « non-corail »).

#### **3.4.2.1 La présegmentation non supervisée**

L'algorithme de segmentation automatisé utilise deux principes importants. D'abord, l'image est découpée en une multitude de segments au cours d'une procédure de présegmentation non supervisée. Ce processus découpe l'image en plusieurs segments contigus dont les frontières reposent sur les discontinuités observées. Finalement, un algorithme intelligent tente d'identifier si ces segments appartiennent au sujet de l'image (au corail) ou non.

Pour effectuer l'étape de présegmentation, d'innombrables méthodes sont disponibles. Nous désirons toutefois une méthode qui respecte deux critères importants : les segments générés doivent être fermés et contigus; et leurs frontières doivent reposer sur les discontinuités de l'image qui permettent bien de départager le « corail » du « non-corail ». Suite à plusieurs tests, deux méthodes simples issues de la théorie de la VA ont retenu notre attention : le LoG (laplacien du gaussien) tel que décrit par Gonzalez et Woods (2008) et le partage des eaux (*watershed*) tel que décrit par Meyer (1994). Ces méthodes sont expliquées en détail à la section 1.3.3. La méthode LoG assure une détection des discontinuités pertinentes d'une image, mais fournit simplement un ensemble discontinu des pixels qui appartiennent aux diverses frontières. Par conséquent, les frontières résultantes du LoG ne sont pas fermées et ne forment pas de segments d'images utilisables.

Inversement, le partage des eaux résulte toujours en de multiples frontières continues et fermées. Cet algorithme forme donc des segments exploitables. En revanche, le partage des eaux est très sensible et la segmentation qui en résulte admet plusieurs milliers de segments de minuscule taille. La solution proposée est une hybridation de ces deux méthodes, ce qui permet de soutirer uniquement leurs avantages. Il suffit d'appliquer l'algorithme de partage des eaux sur l'inverse de la représentation des gradients LoG de l'image pour obtenir de grands segments contigus dont les frontières sont fortement liées aux discontinuités observées.

Initialement, nous considérons cette hybridation comme une idée originale. Toutefois, trop simple pour ne pas avoir déjà été utilisée, de nombreuses publications traitent de principes similaires, par exemple les travaux de Vanhamel, Pratikakis et Sahli (2003). Considérant cette transformée hybride, la Figure 3.11 présente les segments images extraits sur une image de *Coral ID* (les frontières des segments sont en blanc).

Notez que contrairement à l'outil de segmentation manuelle, pour extraire les frontières, l'espace colorimétrique Lab (McLaren, 1976) est utilisé. Dans l'espace Lab, l'intensité lumineuse est complètement indépendante du chromatisme. De ce fait, l'espace Lab permet

une détection plus stable des frontières qu'avec l'espace RVB. L'outil de segmentation manuelle a toutefois été implémenté avec l'espace RVB puisque de cette façon il se comporte de façon prévisible pour l'utilisateur puisqu'il est identique aux outils connus.

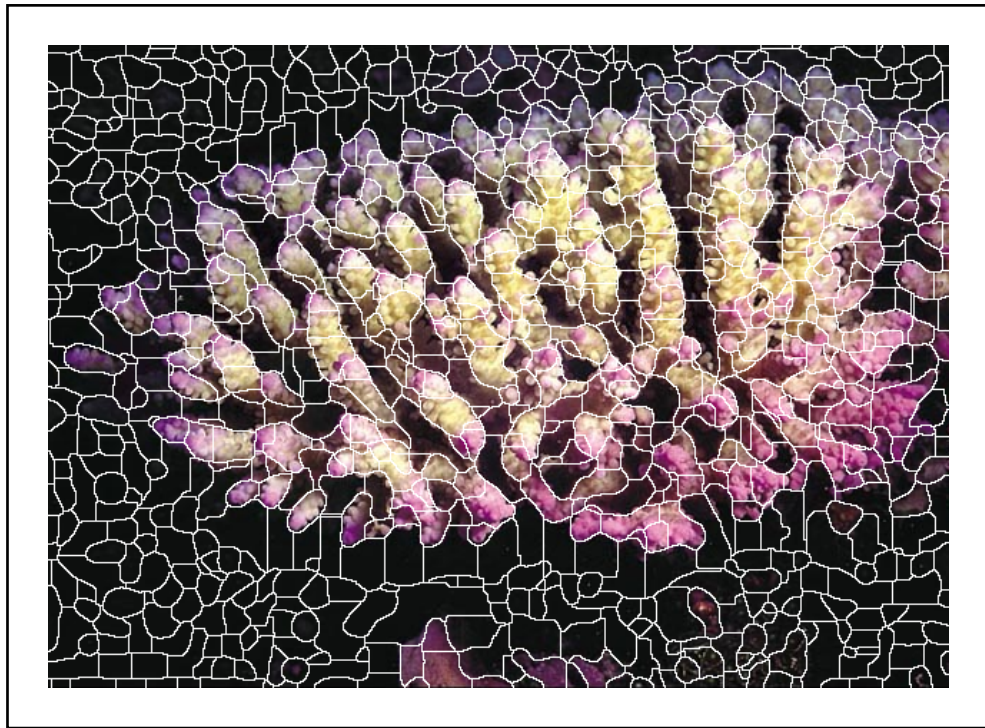


Figure 3.11 Image segmentée par partage des eaux sur le gradient LoG.  
Image originale tirée de *Coral ID* (Veron et Stafford-Smith, 2002)

Notez également que deux paramètres doivent être définis pour permettre cette segmentation automatisée : la taille du filtre LoG et l'écart-type de sa gaussienne. Les paramètres ont été déterminés empiriquement et manuellement à l'aide d'un petit ensemble d'images tirées de *Coral ID* (les cinq premières de notre échantillonnage). Dans un monde idéal, ce type de paramétrisation devrait être optimisé sur un ensemble représentatif de données. Toutefois, notre ensemble est limité et un ajustement manuel des paramètres est suffisant pour faire la preuve de concept. La taille de filtre qui nous semble optimale, suite à une observation qualitative des résultats sur l'ensemble de test, est de 51 pixels de côté et la gaussienne utilisée possède un écart-type de  $51/8$ , soit 6,375 pixels. Cette combinaison a été sélectionnée en s'assurant que la forme du filtre LoG est bien échantillonnée et respecte le « chapeau mexicain » généralement observé qui garantit un résultat pertinent (voir Figure 3.12).

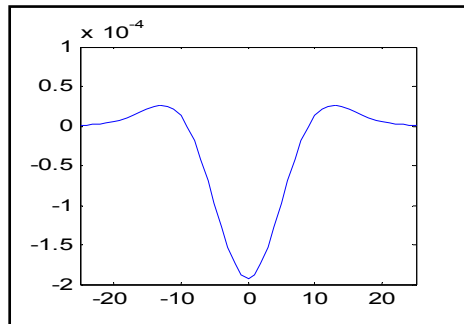


Figure 3.12 Tranche centrale du filtre LoG bidimensionnel (taille 51,  $\sigma = 51/8$ ).

Maintenant que nous avons présélectionné les images, nous utilisons une autre heuristique pour faire l'amalgame de divers segments jugés semblables ou trop petits pour être utilisables. L'algorithme de présegmentation basé sur le partage des eaux appliqué à la réponse au LoG n'est pas parfait. Il génère parfois trop de frontières et par conséquent de nombreux petits segments qui possèdent une surface de quelques pixels seulement. Les artefacts de la compression JPEG des images ont tendance à perturber les canaux couleur de la représentation Lab, ce qui génère une partie de ces frontières indésirables. Puisque nous allons par la suite extraire un lot de descripteurs sur ces zones pour nous permettre de les classer correctement une taille minimale des segments est requise.

Pour éviter ces problèmes, deux correctifs ont été apportés soit - 1 - joindre les segments d'intensité lumineuse moyenne similaire et - 2 - joindre les segments de petite taille à leur voisin immédiat le plus près, en termes d'intensité lumineuse moyenne. La Figure 3.13 illustre la séquence d'opérations nécessaires pour accomplir ces tâches. Suite au traitement, plusieurs zones ont été fusionnées. La Figure 3.14 illustre le résultat. Ce mécanisme a fait passer de 945 à 194 le nombre de segments à traiter dans l'image. De ce fait, l'aire moyenne des segments est passée de près de 300 pixels à plus de 1450 pixels. La taille accrue des zones extraites garantira une stabilité et une représentativité des descripteurs de la morphologie, de la texture et de la couleur. Ces descripteurs seront ensuite utilisés pour généraliser sur l'appartenance d'un segment à la classe « corail » ou « non-corail ».



**Entrées :**  $lb$  : matrice des libellés des segments,  $L$  : matrice d'intensité lumineuse  
 $n$  : nombre de segments dans  $lb$

1. calculer l'intensité moyenne des régions  $im(lb) = \frac{\sum(L \in lb)}{|L \in lb|} \forall lb$
2. pour toute intensité  $i \in \{0,2; 0,4; \dots; 1,0\}$ 
  3. trouver l'ensemble des régions d'intensité similaire  $lb_i \in (i - 0,2) > im_{lb} \geq i$
  4. fusionner les régions similaires  $lb_i$  en exécutant  $lb(lb_i) = n + 1 ; n = n + 1$
  5. relibeller les régions contigües et mettre à jour  $lb, n$  et  $im(lb)$
7. calculer les dimensions en  $x$  et en  $y$  de la boîte  $b$  englobant les segments  $lb$   
 $b(lb, d) \mid lb \in \{1, \dots, n\} ; d \in \{x, y\}$
8. pour toute les petites régions  $lb_p \mid b(lb, d) \leq 5$  pixels de coté
  9. trouver les régions voisines  $lb_v$  de la région  $lb_p$  courante
  10. fusionner la région  $lb_p$  courante avec la région voisine la plus similaire  $lb_s$   
 $lb(lb_p) = lb_s \mid lb_s = \min_{lb} \left( \text{abs} \left( mi(lb_v) - mi(lb_p) \right) \right)$
  11. relibeller les régions contigües et mettre à jour  $lb, n$

Figure 3.13 Algorithmes heuristiques de fusion des segments.

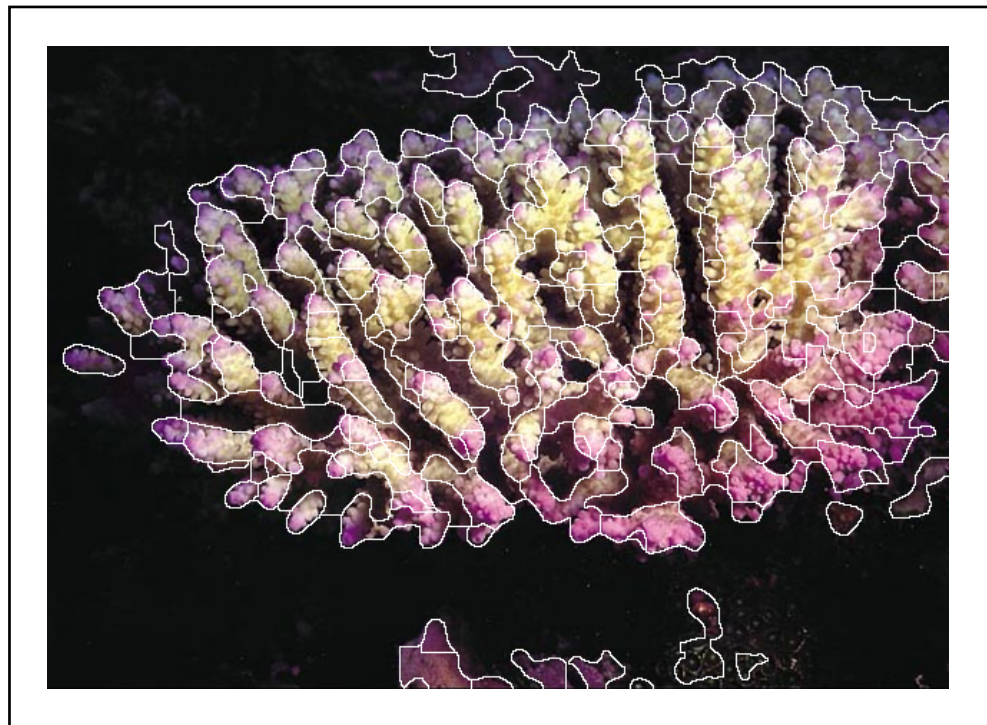


Figure 3.14 Image segmentée après fusion des segments.  
 Image originale tirée de *Coral ID* (Veron et Stafford-Smith, 2002)

### 3.4.2.2 L'extraction des caractéristiques pour la segmentation

Maintenant que la présegmentation non supervisée est complétée, l'heure est venue d'entamer la portion RDF de la problématique de segmentation. Notez que d'un point de vue général, la segmentation automatisée des spécimens de corail est simplement une étape préalable au problème plus vaste et plus ardu de la reconnaissance de l'espèce au niveau taxinomique. De ce fait, plusieurs des étapes sous-jacentes à l'extraction des descripteurs (nommés plus généralement « caractéristiques » en RDF), à l'entraînement d'un algorithme de classification et à l'évaluation des performances de ceux-ci sont communes à l'étape de segmentation et à l'étape d'identification (classification) de l'espèce. Pour alléger la présentation, nous n'allons que présenter les grandes lignes des étapes nécessaires à l'entraînement d'un système de détection de corail pour l'instant. Pour plus de détails, des références seront faites aux sections pertinentes des parties 3.5, 3.6 et 3.7 du document.

À partir de ce point, nous considérons les segments d'images obtenus en 3.4.2 comme des éléments d'un problème de reconnaissance de forme qui doivent être classifiés. Nous avons choisi de considérer un segment d'image comme appartenant véritablement au corail si plus de 50 % de sa superficie se retrouve dans la zone segmentée manuellement en 3.4.1. Pour chacune de ces zones, un lot de 44 descripteurs ont été extraits et consignés dans un fichier au format ARFF. Tel que mentionné précédemment, ce format de fichier est compatible avec la boîte à outils Weka (Hall *et al.*, 2009) et ouvre la porte à tout un monde d'algorithmes de classification et de traitement des données.

L'extraction des caractéristiques des segments d'images a été implémentée en s'inspirant des travaux de Levasseur (2008). Ce lot de descripteurs caractérise la texture, la couleur, la forme et la position des segments de corail présents dans l'image segmentée. Notez que nous avons considéré l'information de position, puisque dans notre banque d'image, le sujet est pratiquement toujours au centre. Dans un contexte réel, il sera important d'optimiser l'ensemble des descripteurs pour l'adapter davantage aux conditions présentes. Le nouvel ensemble de description devra alors exclure la position relative du segment, puisqu'elle ne

sera plus pertinente (à moins que le nouveau protocole de photographie centre également les spécimens dans l'image). Ce genre de détails peut faire une grande différence dans les performances observées. Notez que les mêmes caractéristiques ont été utilisées pour effectuer la tâche de segmentation et de classification des spécimens. Par conséquent, nous avons évité de répéter les informations sur cette extraction de caractéristiques à deux reprises. Pour tous les détails sur cette extraction de caractéristiques, la section 3.5 saura vous informer.

### 3.4.2.3 L'entraînement d'un algorithme de RDF pour la segmentation

Une fois les caractéristiques extraites, nous ne travaillons plus avec les images directement. Seule une table de données nous reste (le fichier ARFF), avec sur chaque ligne l'information relative à un segment d'une image, et où chaque colonne détaille le spécimen étudié par les caractéristiques numériques extraites précédemment. C'est à partir de cette table de nombres que l'algorithme de RDF est entraîné. Pour toute cette problématique, nous avons choisi d'utiliser le SVM (séparateur à vaste marge), tel que décrit par Cortes et Vapnik (1995). Tout comme pour l'extraction des caractéristiques, nous avons utilisé le même protocole expérimental qu'en classification. Le choix du SVM et de l'implémentation choisie sont détaillés en 3.6. De plus, les résultats du SVM seront comparés avec ceux du kPPV, avec  $k = 1$  voisin. La seule particularité de l'implémentation pour la segmentation automatisée est la classe de l'élément. Pour entraîner les algorithmes de RDF, les caractéristiques de plusieurs exemples de segments de *corail* et de *non-corail* doivent être présentées à l'algorithme. Ces exemples doivent être accompagnés d'une étiquette qui informe l'intelligence artificielle de leur véritable nature (*corail* ou *non-corail*) pendant la phase d'apprentissage. La nature d'un élément à classifier est nommée la « classe » dans la terminologie de la reconnaissance de forme. Il va de soi que la classe utilisée pour la problématique de segmentation est différente de celle utilisée pour l'identification.

Enfin, les résultats obtenus suite à la mise à l'essai de l'algorithme de segmentation automatisée seront présentés à la section 3.8.

### 3.5 L'extraction des caractéristiques pour la classification

Cette section traite de l'extraction de descripteurs à partir des segments d'images qui ont été extraits par la procédure de présegmentation de 3.4.2.1. Les informations présentées ici sont valables autant pour l'entraînement d'une IA de segmentation que pour la classification taxinomique des spécimens.

Dans un problème de RDF, l'extraction de caractéristiques permet de simplifier les données en les transposant dans un espace de description de taille réduite. La taille de l'espace de description est très importante. Plus l'espace est de grande dimension, plus les échantillons considérés sont éloignés les uns des autres. Il devient alors difficile (et couteux) de découvrir (ou de modéliser) les regroupements naturels des données (modes), lorsque l'on considère un nombre fini d'échantillons. Tel que mentionné précédemment, ce phénomène est appelé la *malédiction de la dimensionnalité* en statistique (Bellman, 1961). En plus de la réduire, l'extraction de caractéristiques permet d'uniformiser la dimension des vecteurs échantillons considérés. Dans notre problématique chaque échantillon est un segment d'image de taille variable où l'élément descriptif le plus simple est le *pixel*. En moyenne, l'espace de représentation initial s'exprime en 4 350 (1 450 pixels  $\times$  3 couches couleurs) dimensions tel que présenté en 3.4.2.1. En exécutant l'extraction de caractéristiques sur chaque segment, cet espace est réduit et uniformisé à un vecteur de taille constante à 44 dimensions.

Les descripteurs que nous avons choisis sont en partie inspirés des travaux de Levasseur (2008). Nous avons choisi de s'inspirer de ces travaux pour plusieurs raisons. Premièrement, les travaux de Levasseur démontrent que ces descripteurs sont suffisamment discriminants pour résoudre plusieurs problèmes de classification d'objets biologiques à partir d'images numériques. En effet, s'il était possible d'obtenir une performance en classification similaire à celle obtenue dans les travaux de Levasseur pour résoudre le problème d'identification des coraux, ce serait amplement satisfaisant pour notre application. Avec de telles performances, le système d'identification de corail serait un outil indispensable pour la communauté scientifique qui s'intéresse au suivi de la biodiversité.

De plus, les descripteurs ciblés par Levasseur appartiennent à un groupe de statistiques fondamentales largement utilisées. De ce fait, elles sont bien documentées et on retrouve leur implémentation dans plusieurs boîtes d'outils de VA telles qu'OpenCV (Bradski, 2000), ImageJ (Abramoff, Magelhaes et Ram, 2004), MATLAB®, ou ImageMagick (Still, 2006). Au cours de cette section chaque descripteur sera présenté en détail et des justifications seront apportées quant à leur utilisation dans ce contexte.

Notez que pour la problématique de l'identification taxinomique, nous ne considérons uniquement que les segments d'images qui appartiennent véritablement au sujet photographique (au corail) de chaque image. Comme mentionné à la section précédente, nous avons choisi de considérer un segment d'image comme appartenant au corail si plus de 50 % de sa superficie se retrouve dans la zone segmentée manuellement en 3.4.1. Encore une fois, pour chaque zone, le lot de 44 descripteurs a été extrait et a été consigné dans un fichier au format ARFF. Finalement, dans le contexte d'identification, l'étiquette de classe attribuée à chaque échantillon décrit au niveau taxinomique le segment d'image.

Les 44 descripteurs sont décrits par les sous-sections suivantes. La présentation de la théorie derrière chaque calcul est partagée entre cette section et la section correspondante de la revue de la littérature. Des références à la section 1.3.4 seront mentionnées, au besoin. Pour chaque descripteur, la cardinalité est exprimée entre parenthèses. Notez que l'implémentation à cette étape a été effectuée à l'aide de MATLAB® et de son « *Image Processing Toolbox* ». Les formules mathématiques et descriptions qui seront présentées dans les sections suivantes sont donc inspirées de la documentation de cet outil et du mémoire de Levasseur (2008).

### **3.5.1 Les descripteurs de morphologie**

Les descripteurs de morphologie décrivent la forme des segments d'images issus de la présegmentation automatisée. Ces sept descripteurs sont extraits comme suis :

**Le centroïde du segment (2)**

Le centroïde est un concept physique qui a été porté à la géométrie de la forme pixelisée du segment. Il s'exprime par la coordonnée en  $x$  et en  $y$  du pixel au centre de masse de la région.

**L'aire du segment**

L'aire du segment est simplement la quantité de pixels comprise dans la zone étudiée.

**L'excentricité du segment**

L'excentricité est un nombre scalaire qui exprime l'excentricité de l'ellipse qui possède le même moment d'inertie que la forme étudiée. L'excentricité de l'ellipse est alors calculée par le ratio de distance entre le foyer de l'ellipse et la longueur de son axe majeur. Une excentricité de 0 est obtenue pour un cercle alors qu'une excentricité de 1 est obtenue pour un segment de droite.

**La taille de l'axe majeur et mineur (2)**

La longueur en pixels de l'axe majeur et mineur d'une ellipse ayant le même moment d'inertie que la forme en question.

**L'orientation de l'axe majeur**

L'orientation en degrés de l'axe majeur d'une ellipse ayant le même moment d'inertie que la forme en question, par rapport à l'axe des  $x$  de l'image. ( $\pm 90^\circ$ )

**3.5.2 Les descripteurs de texture**

Les descripteurs de texture utilisent diverses statistiques sur l'histogramme des tons de gris. Notez que plusieurs de ces statistiques sont élémentaires et ne seront pas détaillées.

**La moyenne d'intensité lumineuse**

Voir l'équation (1.5) de la section 1.3.4.1.

**L'écart type de la distribution des tons de gris**

Voir l'équation (1.6) de la section 1.3.4.1.

**La valeur R du second moment de l'histogramme des tons de gris**

Le second moment  $m_2$  est la variance de population évaluée sur l'histogramme des tons de gris. La valeur R se calcule quant à elle comme illustré à l'équation (1.7) de la section 1.3.4.1.

**Le coefficient d'asymétrie de la distribution des tons de gris**

Se calcule comme illustré à l'équation (1.8) de la section 1.3.4.1 à partir de l'histogramme des tons.

**Le coefficient d'aplatissement de la distribution des tons de gris**

Se calcule comme illustré en (1.9) de la section 1.3.4.1 à partir de l'histogramme.

**Le coefficient d'uniformité de la distribution des tons de gris**

Se calcule à partir de l'histogramme des tons, voir l'équation (1.10) de la section 1.3.4.1.

**L'entropie de la distribution des tons de gris**

L'entropie calculée sur l'histogramme des tons à partir de l'équation (1.11) de la section 1.3.4.1.

**Contraste à partir de la matrice de cooccurrence des tons de gris (3)**

Pour toutes les caractéristiques basées sur la matrice de cooccurrence des tons de gris, la section 1.3.4.2 vous guidera sur le sujet. Cette caractéristique est calculée horizontalement, verticalement et diagonalement, d'où sa cardinalité de trois.

**Corrélation à partir de la matrice de cooccurrence des tons de gris (3)**

Voir 1.3.4.2.

**Énergie à partir de la matrice de cooccurrence des tons de gris (3)**

Voir 1.3.4.2.

**Homogénéité à partir de la matrice de cooccurrence des tons de gris (3)**

Voir 1.3.4.2.

**3.5.3 Les descripteurs de couleur**

Les descripteurs de couleur sont de simples statistiques extraites sur chaque canal couleur d'un segment.

**Mode du canal couleur (3)**

Mode statistique de la distribution de l'intensité de chaque canal couleur.

**Moyenne du canal couleur (3)**

Moyenne d'intensité de chaque canal couleur, idem à l'équation (1.5) de la section 1.3.4.1.

**Maximum du canal couleur (3)**

Valeur d'intensité maximale que prends chaque canal couleur.

**Minimum du canal couleur (3)**

Valeur d'intensité minimale que prends chaque canal couleur.

**Somme d'un détecteur de discontinuité de Sobel du canal couleur (3)**

Pour obtenir cette caractéristique il suffit d'effectuer la somme des pixels qui répondent au filtre de Sobel sur chaque canal couleur, voir section 1.2.1 et 1.3.3.1.

**Entropie du canal couleur (3)**

L'entropie calculée sur l'histogramme des tons de chaque couche couleur  $P_i$  et est obtenu à partir de l'équation (1.11) de la section 1.3.4.1.



### 3.5.4 Divergences relatives aux travaux de Levasseur

Notez que la présentation des caractéristiques des sections précédentes n'inclut pas tous les descripteurs suggérés par Levasseur (2008). Nous avons choisi d'éliminer certains d'entre eux, comme ceux relatifs à la transformée de Fourier de l'image et ceux relatifs à l'étude des LBP (*local binary patterns*) pour plusieurs raisons. Une erreur conceptuelle et d'implémentation au sujet des caractéristiques basées sur Fourier se solde par un lot de descripteurs non discriminants. Une étude plus poussée est nécessaire à ce sujet, et elle est hors du cadre théorique qui guide cette recherche. Du côté des LBP, en plus de leur coût computationnel élevé, il semble que ces descripteurs soient très sensibles aux variations locales de la texture et ne soient pas très pertinents dans le contexte de segmentation et de classification automatisée des images de corail.

La performance de discrimination de ces deux types de caractéristiques de texture sur notre problématique a été vérifiée rapidement dans l'environnement Weka (Hall *et al.*, 2009). Diverses méthodes d'évaluation (*ranking*) ont validé nos hypothèses à ce sujet : les caractéristiques relatives à Fourier et aux LBP n'ont pas prouvé leur capacité à départager le corail du non-corail, ou encore à déterminer l'espèce précise d'un spécimen.

### 3.6 L'entraînement des classificateurs et l'optimisation des métaparamètres

À présent, nous avons traité de la taxinomie des coraux, nous avons échantillonné un ensemble de données pour conduire notre étude, nous avons segmenté les spécimens de corail au sein des images et finalement, nous avons réduit le problème de classification à une modélisation dans un espace à 44 dimensions, par le biais de l'extraction de caractéristiques inspirée en partie de Levasseur (2008). Au cours de cette section, nous compléterons le travail, en entraînant deux algorithmes de RDF à reconnaître les spécimens de corail au niveau taxinomique.

Notez que la procédure d'entraînement présentée ici est exactement la même que celle qui est utilisée pour l'intelligence de segmentation automatisée, à la seule différence près de la classe attribuée à chaque échantillon. Pour le problème de classification, la classe est de nature taxinomique. Chaque échantillon est donc identifié par l'union de l'information de la famille, du genre et de l'espèce. Dans le cas de la segmentation, chaque échantillon est identifié par l'attribut « corail » ou « non-corail ». Notez toutefois que comme il est justifié au cours de la section 3.4.2 le concept de « corail » et de « non-corail » n'est pas tout à fait compatible avec le nom choisi pour ces deux classes. La classe « corail » est attribuée aux segments d'images qui appartiennent au sujet photographique. La classe « non-corail » contient l'arrière-plan photographique, et par conséquent, elle contient parfois des segments de coraux qui sont à l'arrière-plan dont la nature taxinomique nous est inconnue.

L'algorithme de RDF que nous avons sélectionné pour cette étude est le SVM (*support vector machine* – séparateur à vaste marge), tel que décrit par Cortes et Vapnik (1995). Dans notre étude précédente, au CHAPITRE 2, le PM (perceptron multicouches) a toujours semblé l'algorithme le plus performant du lot, surpassant le SVM. Toutefois, au cours de diverses tentatives pour améliorer les performances du système, nous avons découvert l'importance de l'optimisation des métaparamètres. Après une procédure d'optimisation simpliste, le SVM surpasse largement les performances du SVM non-optimisé, et également celles du PM non-optimisé.

Le SVM possède relativement peu de métaparamètres, contrairement au PM dont l'entière topologie peut être remodelée en une infinité de possibilités. L'optimisation du PM est donc très fastidieuse et nécessite une fine connaissance du domaine de l'optimisation et de ses méthodes stochastiques. Ces méthodes permettent de découvrir une solution optimale dans un espace à très grande dimension, tel qu'occasionné par la grande flexibilité du PM et ce, en un temps relativement court. Toutefois, toutes ces théories d'optimisation sont un vaste sujet de recherche en soit, et sont hors du cadre théorique qui guide ce travail.

Pour limiter la complexité de l'approche d'entraînement, nous avons donc sélectionné le SVM, parmi le lot de méthodes de RDF de la littérature. De plus, comme pour le PM, les résultats du SVM non-optimisé se sont démarqués des autres algorithmes que nous avons testés au CHAPITRE 2, ce choix est donc autant appuyé par la simplicité que la performance de l'algorithme. L'implémentation que nous avons choisie pour entraîner et mettre à l'épreuve le modèle est LIBSVM de Chang et Lin (2001) puisqu'elle est très bien documentée et très populaire dans le domaine.

Notez également que pour des fins de comparaison, l'algorithme du plus proche voisin sera aussi mis à l'épreuve. Nous pourrions ainsi vérifier si un algorithme de classification complexe, tel le SVM apporte un gain significatif dans cette problématique.

### **3.6.1 L'optimisation des métaparamètres**

Pour optimiser les métaparamètres du SVM, une procédure détaillée est suggérée par les auteurs de LIBSVM (Hsu, Chang et Lin, 2003). Cette procédure fournit une méthode d'optimisation pour certains paramètres et suggère également l'utilisation de valeurs généralement recommandées pour d'autres. Au cours de nos expérimentations, nous n'avons aucunement divergé de cette procédure. Les prochaines pages résumeront donc les suggestions apportées par ce document.

En premier lieu, notez que LIBSVM implémente plusieurs versions de l'algorithme du SVM. La version nommée C-SVC sera utilisée et correspond à la définition initiale du SVM tel qu'introduite par Cortes et Vapnik (1995). Le second paramètre fondamental est le choix du type de noyau. Pour arriver à séparer facilement les classes dans un espace donné, le SVM utilise une projection dans un espace augmenté en utilisant la notion de noyau (voir la section 1.3.5.5 pour plus de détails). Pour ce faire, le noyau RBF (*radial basis function* – fonction à base radiale) a été sélectionné. Ce noyau permet de gérer les cas où la relation entre les caractéristiques et les classes n'est pas linéaire. De plus, le choix de ce noyau implique un faible nombre de métaparamètres, ce qui simplifie la recherche de la configuration optimale.

Deux autres paramètres statiques (non-optimisés) sont suggérés également et sont relatifs uniquement à l'implémentation spécifique de LIBSVM. Premièrement, l'algorithme implémenté par LIBSVM inclut une heuristique d'arrêt hâtif de l'entraînement. Cette heuristique permet d'économiser les ressources en termes de temps de calcul. Le guide nous suggère de l'activer à prime abord. Enfin, le dernier critère suggéré est la taille du cache. Pour accélérer les calculs, l'implémentation du C-SVC conserve en mémoire (en cache) certaines données, la taille du cache d'apprentissage suggérée est de 100 Mo (méga-octets).

Enfin, les deux paramètres restants ont un impact majeur sur les performances et doivent être optimisés selon le contexte précis de l'apprentissage. Pour ce faire, le guide suggère une recherche de la configuration optimale en essayant de multiples combinaisons de paramètres. Les deux paramètres à optimiser sont le coût de l'erreur  $C$  et la variance du noyau RBF  $\gamma$  (gamma). Étant donné le faible nombre de paramètres, et la nécessité de couvrir l'espace des possibilités, la recherche sera effectuée dans une grille (*grid search*).

Pour ce faire nous allons explorer l'espace des paramètres comme suis. Pour le paramètre  $C$  les valeurs testées sont toutes les puissances impaires de 2 entre  $2^{-5}$  et  $2^{15}$ . Du côté de  $\gamma$  toutes les puissances impaires de 2 seront testées entre  $2^{-15}$  et  $2^5$ . Au total, le produit scalaire de ces deux espaces donne naissance à 121 configurations possibles ( $11 \times 11$ ). Pour chaque configuration, un SVM est entraîné sur la portion de validation de la validation croisée (1/8 de la base de données) tel que détaillé en 3.3.2. Cet entraînement est effectué au sein d'une autre validation croisée à 5 partitions, ce qui permet d'assurer la stabilité des résultats obtenus. Après avoir entraîné les 121 classificateurs sur autant de configurations possibles, le meilleur résultat en termes de taux de classification correcte est retenu.

À partir du moment où l'optimum est obtenu sur cette grille grossière de paramètres, une deuxième itération est exécutée pour affiner le résultat. À cette étape le voisinage de l'optimum  $O^p$  est subdivisé à nouveau en 121 configurations avec 11 puissances distribuées uniformément entre  $O^{p-2}$  à  $O^{p+2}$  pour chacun des deux paramètres. Suite à l'entraînement en validation croisée des classificateurs, la paramétrisation optimale pour le SVM est obtenue et

conservée pour la phase d'entraînement. La Figure 3.15 illustre le résultat de l'optimisation par grille pour le premier réplica de l'expérimentation de segmentation automatisée.

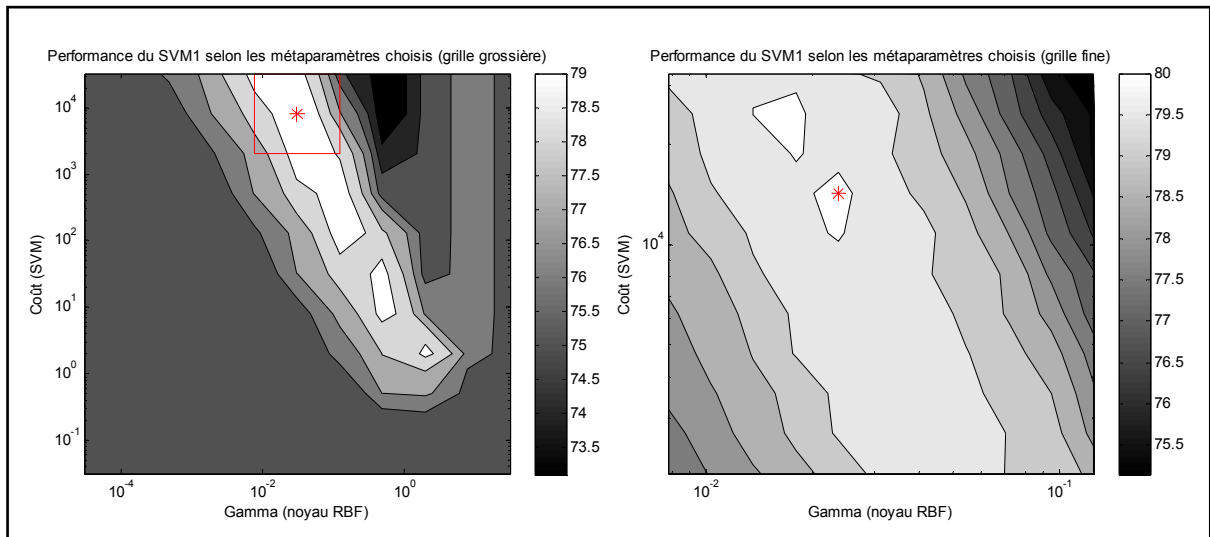


Figure 3.15 Topologie de l'espace d'optimisation pour le SVM de segmentation.

### 3.6.2 L'entraînement des algorithmes de RDF

Au moment d'entraîner l'algorithme, la majeure partie du travail est déjà accomplie. Pour entraîner le SVM, les paramètres optimaux obtenus précédemment sont utilisés sur la portion d'entraînement de la validation croisée (6/8 de la base de données). Une fois l'entraînement terminé, le modèle SVM obtenu est stocké pour une réutilisation ultérieure. Il ne reste plus qu'à tester ce modèle sur des données différentes pour obtenir un estimé de la performance qu'est capable de rendre ce système. Pour plus de renseignements sur la procédure d'apprentissage du SVM, consultez la section pertinente de la revue de littérature en 1.3.5.5. Notez que pour le kPPV, la procédure d'entraînement se limite à mémoriser la base d'apprentissage.

### **3.7 La généralisation**

Maintenant que nous avons entraîné un algorithme de RDF sur le problème de segmentation et/ou de classification d'images de corail, voici comment nous allons évaluer les performances en généralisation du système. Pour commencer, tous les tests ont été exécutés sur une partition de données indépendante de l'optimisation des paramètres et de l'entraînement du classificateur. Cette portion indépendante restante des données (1/8 de la base de données) est utilisée pour accomplir cette tâche.

Au sein de la validation croisée à huit blocs décrite en 3.3.2, le processus d'optimisation, d'entraînement et de test est exécuté huit fois. Les huit réplifications génèrent huit taux de classification des spécimens, tels qu'évalués sur la portion de tests. Ces huit taux sont ensuite modélisés par une distribution de Student qui résulte en un intervalle de confiance sur la performance. Comme mentionné précédemment, étant donné la faible variance observée sur les huit résultats, huit réplifications sont amplement suffisantes pour permettre une comparaison statistique de la performance des systèmes. La méthode de Student est détaillée par la revue de littérature en 1.3.6.4.

### **3.8 Les résultats expérimentaux et leur interprétation**

Pour faire suite à la méthodologie expérimentale, diverses expériences ont été conduites pour valider ou infirmer les hypothèses de travail. Ces expériences traitent des deux sujets, soit la segmentation du sujet photographique d'une image de corail encyclopédique et la classification des segments d'images de corail dans ce contexte particulier. Cette section présente les résultats bruts de chaque expérimentation et les statistiques descriptives liées aux échantillonnages de données utilisées pour la validation, l'apprentissage et les tests.

Notez que deux algorithmes de RDF ont été mis à l'épreuve au cours de ces tests, le SVM et le kPPV avec  $k = 1$  voisin. Le kPPV a été testé uniquement pour valider la configuration des métaparamètres du SVM. Considérant une paramétrisation adéquate, le SVM est généralement plus performant que l'algorithme du kPPV puisqu'il permet de modéliser une

relation complexe. Il est ainsi évident de détecter un problème de paramétrisation du SVM ou toute autre anomalie. Dans ce contexte, une description statistique élaborée des résultats ne sera conduite que pour le SVM. Notez également que toute analyse sera conduite à l'aide des intervalles de Student sur la moyenne à un niveau de confiance de 95 %.

### **3.8.1 Expérience sur la segmentation automatisée**

Pour la segmentation, une seule expérience a été conduite. Nous avons simplement tenté de classer les segments d'images issus de la présegmentation non supervisée (voir 3.4.2.1) parmi les classes *corail* et *non-corail*, telles qu'elles ont été définies en 3.4.2. Tout d'abord, voici une description du résultat de la présegmentation automatisée. Cette description sera suivie des taux de classification des algorithmes de RDF entraînés à segmenter le corail.

#### **3.8.1.1 Analyse de la présegmentation non supervisée**

À titre de rappel, la présegmentation est responsable de découper chaque image en segments contigus dont les frontières reposent sur les discontinuités observées. Après l'exécution de l'algorithme, les segments formés sont de taille très variée, mais permettent de bien discerner le *corail* du *non-corail*. Lorsque plus de 50 % de la superficie d'un segment se retrouve dans la zone segmentée manuellement en 3.4.1 comme étant un *corail*, ce segment est étiqueté avec cette classe. Suite à cette identification manuelle des segments on retrouve 29 913 segments (75,1 %) parmi la classe *corail* et 9 899 segments (24,9 %) parmi la classe *non-corail*. Cette base de données étiquetée est la référence qui servira à l'entraînement de l'algorithme de RDF de segmentation.

Avant de passer aux résultats de segmentation, voici quelques statistiques sur les segments de classe *corail* issus de la présegmentation non supervisée. En premier lieu, la Figure 3.16 présente un histogramme de la distribution de la taille des segments de *corail*. Cette taille est importante, puisque dans les sections suivantes les segments de corail seront utilisés pour généraliser sur l'espèce du segment.

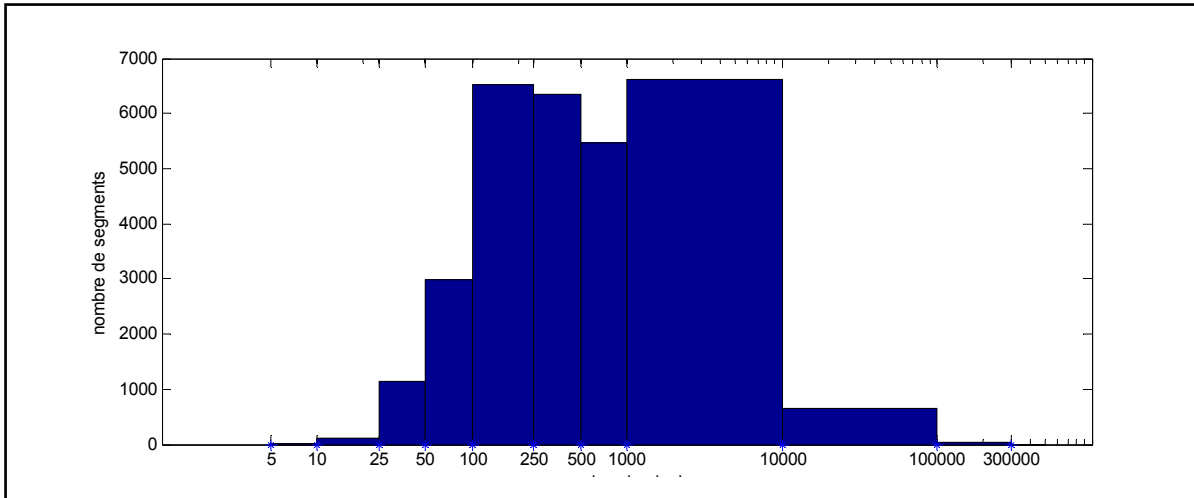


Figure 3.16 Distribution de l'aire des segments de corail issus de la présegmentation.

De plus, la Figure 3.17 présente la distribution du nombre de segments par classe, ce qui indique exactement la quantité de données disponible pour l'entraînement lors de la prochaine expérimentation qui traite de la classification de l'espèce du *corail*. Notez que les nombres présentés en abscisse identifient les classes et sont conséquents avec le Tableau 3.1.

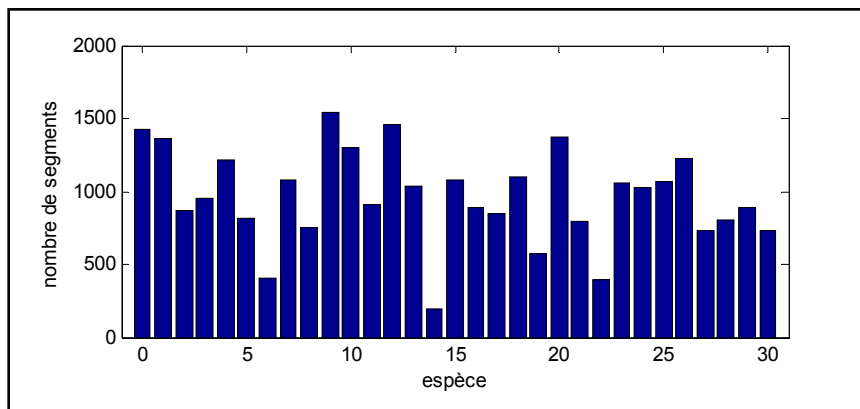


Figure 3.17 Nombre de segments de corail en fonction de leur classe.

La Figure 3.17 informe sur la quantité d'information rendue disponible au système classificateur. Toutefois, le nombre de segments n'est pas directement lié à la quantité d'informations extraite de l'espace image. Pour connaître cette quantité, la Figure 3.18 présente la répartition de l'aire totale (en pixels) des segments de *corail* disponible pour



chaque espèce (classe). Notez que cette superficie est bien évidemment influencée par le nombre d'images disponibles pour chacune des classes (voir Tableau 3.1).

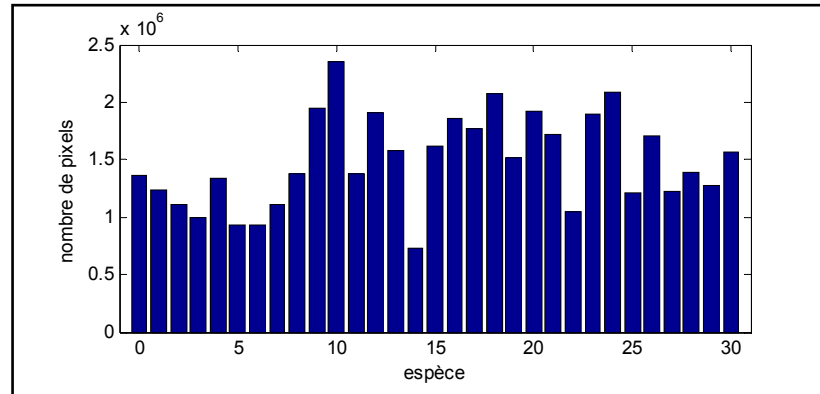


Figure 3.18 Superficie totale des segments de corail en fonction de leur classe.

### 3.8.1.2 Évaluation de la performance en segmentation

Les résultats de segmentation automatisée s'expriment simplement : l'expérience a été conduite sur l'échantillonnage complet des 31 classes et a été formulée sous une validation croisée à huit partitions. Les résultats qui en découlent sont présentés par le Tableau 3.3.

Tableau 3.3 Taux de classification bruts de segmentation

Itération	SVM	kPPV	Itération	SVM	kPPV
1	80,8 %	72,1 %	5	79,9 %	72,1 %
2	79,7 %	71,9 %	6	80,1 %	72,5 %
3	80,2 %	72,1 %	7	80,2 %	71,6 %
4	81,1 %	72,8 %	8	79,8 %	72,9 %

Suite à l'analyse statistique des résultats présentés ci-haut, on remarque que le **taux de classification moyen** espéré pour un tel système de segmentation est de **80,20 ± 0,41 %** pour le SVM (19 fois sur 20, selon l'intervalle de confiance de Student sur la moyenne, voir

1.3.6.4). Ce niveau de performance est acceptable pour un usage destiné à la reconnaissance du corail (Soriano *et al.*, 2001), de plus il est comparable à la littérature (voir section 1.1). On remarque également que l'utilisation d'un algorithme de RDF évolué (tel le SVM) permet un gain significatif en performances de classification par rapport au **kPPV**; qui admet un **taux de classification moyen** de  $72,26 \pm 0,38 \%$ .

Les matrices de confusion totales du SVM et du kPPV (somme des huit itérations) liées à la segmentation des spécimens sont présentées au Tableau 3.4. Lorsque l'on considère individuellement les huit itérations, la **précision** sur la décision du classificateur observée pour le **SVM** (classe *corail*) est de  $80,80 \pm 0,53 \%$ . Cette statistique signifie que  $80,80 \pm 0,53 \%$  des segments classifiés comme *corail* appartiennent réellement à la classe *corail*. De plus, les mêmes matrices de confusion indiquent un taux de **rappel** de  $96,61 \pm 0,52 \%$  pour la classe *corail* avec le **SVM**. Cette statistique indique quant à elle que  $96,61 \pm 0,52 \%$  des segments d'images de la base de données de test étiquetés comme étant un *corail* (classe réelle) ont été identifiés comme étant un *corail* (classe estimée). En d'autres termes, seulement  $\sim 3 \%$  des segments de *corail* véritable ont été classifiés *non-corail*.

Tableau 3.4 Matrices de confusion de la segmentation automatisée

		Classe réelle (SVM)		Classe réelle (kPPV)	
		Corail	Non-corail	Corail	Non-corail
Classe estimée	Corail	28 900	6 869	24 577	5 709
	Non-corail	1 013	3 030	5 336	4 190

Étant donné la nature subjective de la segmentation manuelle effectuée en 3.4.1, nous avons également compilé une évaluation qualitative de la performance en segmentation pour chaque image (contrairement à la compilation par segments effectuée précédemment). Ainsi, une connaissance plus exacte des erreurs de segmentation a pu être extraite des résultats

manuellement. Suite à l'évaluation qualitative, **166** des **196** images de la base de données (**84,7 %**) ont été évaluées comme « bien segmentées ». La Figure 3.19, la Figure 3.20 et la Figure 3.21 montrent quelques exemples de bonne segmentation. Au contraire, la Figure 3.22, la Figure 3.23 et la Figure 3.24, montrent quelques mauvaises segmentations.

La performance évaluée qualitativement sur les images est plus élevée que celle évaluée quantitativement sur les segments. Ce phénomène est intimement lié à la nature qualitative de la segmentation manuelle qui a servi de référence à l'expérience. En effet, les coraux segmentés manuellement admettent un lot de zones d'ombres et de cavités laissant paraître l'arrière-plan. La nature plus précise de la présegmentation non supervisée segmente ces zones d'ombres ou non homogènes au corail. Ainsi, certaines d'entre elles sont classifiées correctement comme appartenant à l'arrière-plan, mais toutefois dénotées comme étant des erreurs de classification lorsque comparées à la segmentation manuelle.

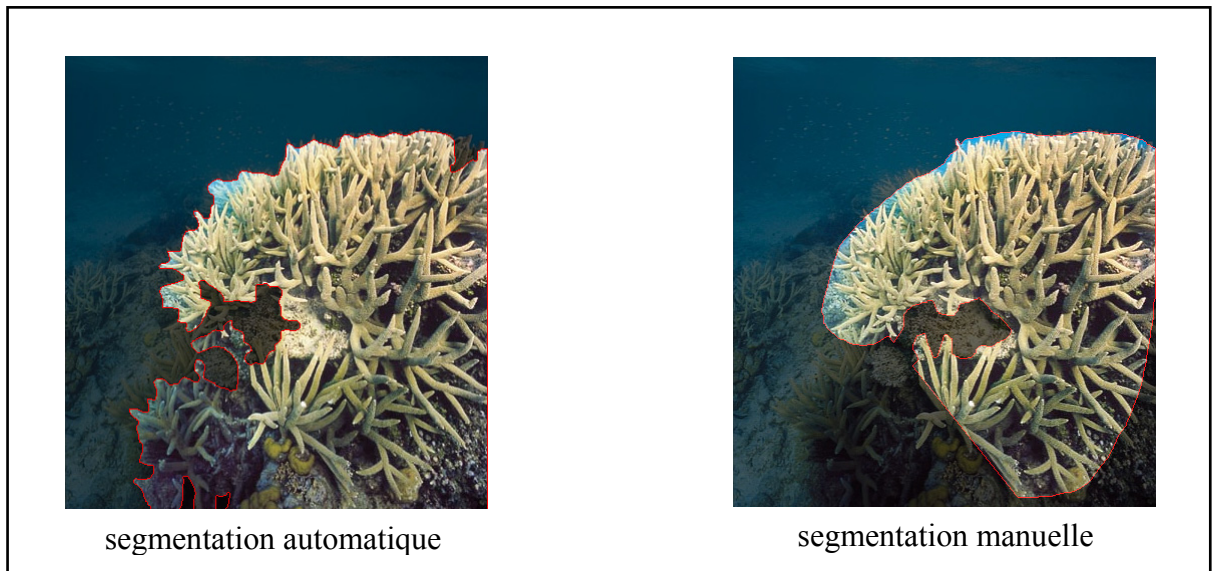


Figure 3.19 Exemple de bonne segmentation (*acropora nobilis* 52-06E).  
Adaptée de Veron et Stafford-Smith (2002)

La Figure 3.19 et la Figure 3.20 démontrent bien la nature plus précise de la segmentation automatisée. Le profil de la segmentation manuelle est plus lâche, moins bien adapté à la complexité morphologique du spécimen de corail.

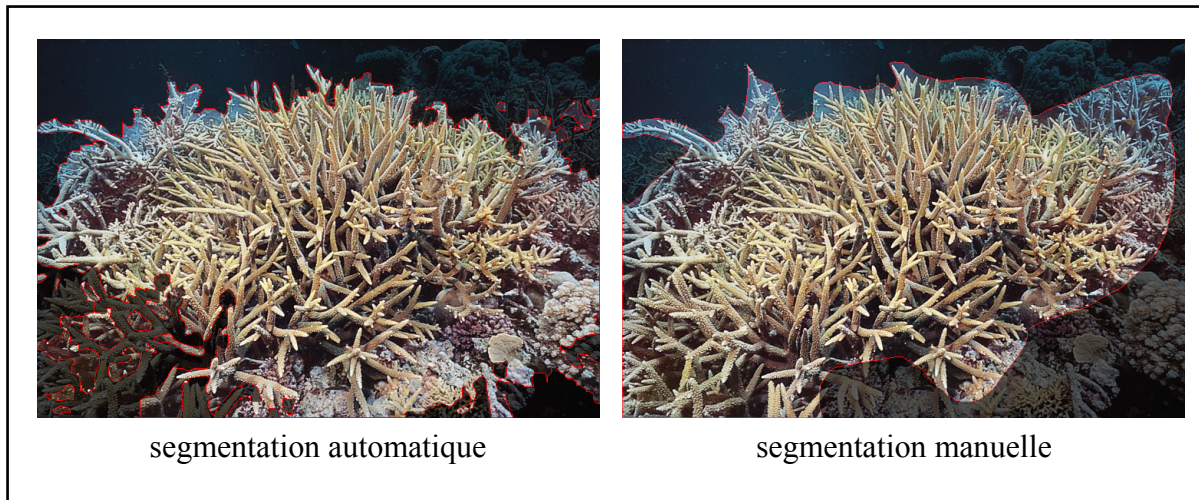


Figure 3.20 Exemple de bonne segmentation (*acropora nobilis* 52-01M).  
Adaptée de Veron et Stafford-Smith (2002)

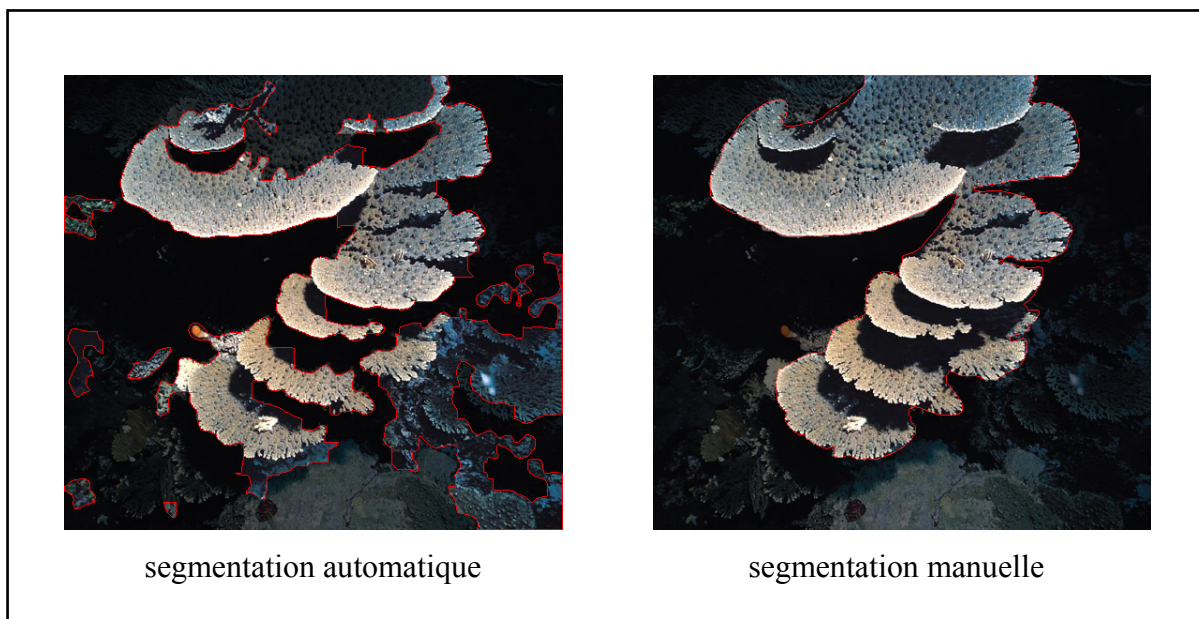


Figure 3.21 Exemple de bonne segmentation (*acropora glauca* 31-01M).  
Adaptée de Veron et Stafford-Smith (2002)

La Figure 3.21 montre la capacité à éliminer les zones d'ombres du spécimen de corail de l'algorithme de segmentation automatisée. Ce comportement représente un avantage pour la classification automatisée ultérieure des segments de corail. Ainsi, les zones ombragées ne contenant aucune information sur la nature du spécimen ne seraient pas considérées.

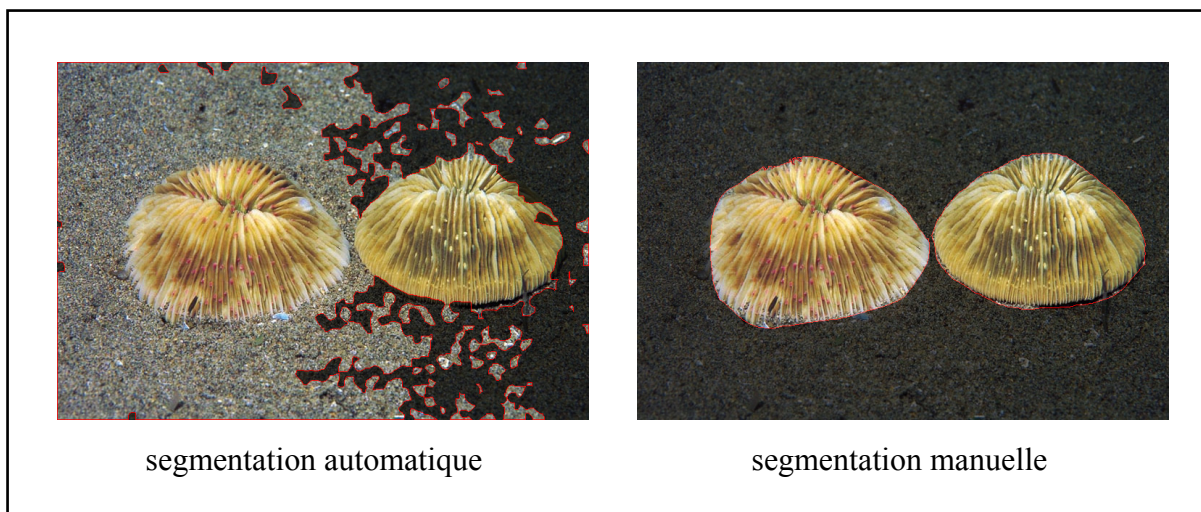


Figure 3.22 Exemple de mauvaise segmentation (*cycloseris cyclolites* 122-09).  
Adaptée de Veron et Stafford-Smith (2002)

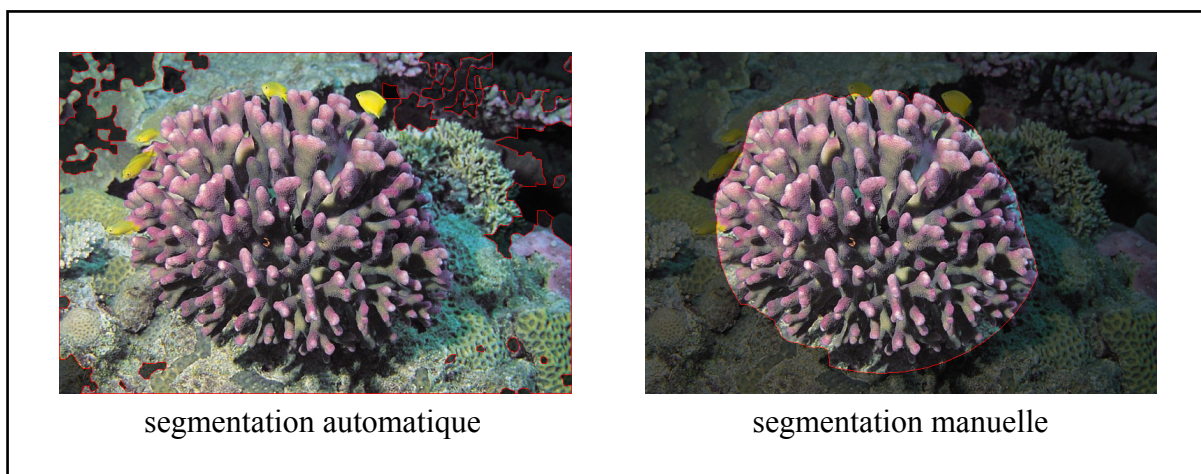


Figure 3.23 Exemple de mauvaise segmentation (*stylophora pistillata* 345-10M).  
Adaptée de Veron et Stafford-Smith (2002)

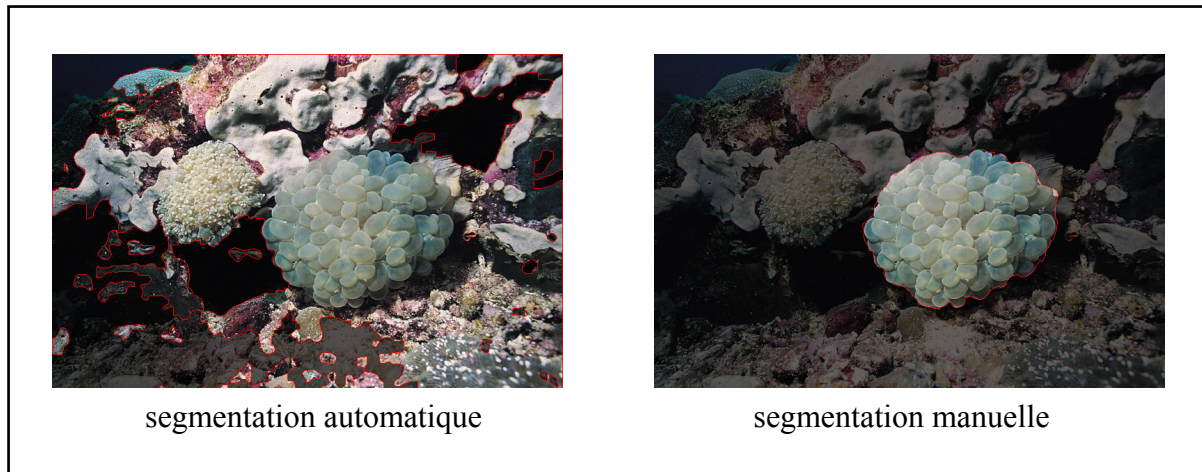


Figure 3.24 Exemple de mauvaise segmentation (*plerogyra sinuosa* 305-07).  
Adaptée de Veron et Stafford-Smith (2002)

Les trois dernières figures démontrent quant à elles quelques exemples d'une segmentation automatisée de mauvaise qualité. Les images comportant un arrière-plan complexe, doté de textures uniformes ou d'autres coraux sont sujettes aux erreurs de segmentation automatisée.

### 3.8.2 Expériences sur la classification automatisée

Au sujet de la classification, plusieurs expériences ont été tentées. En effet, il est difficile de conclure sur la faisabilité d'implémentation d'une IA pour la reconnaissance taxinomique des coraux avec aussi peu de données. Nous avons donc dû varier les expériences pour dresser un portrait plus juste des conditions nécessaires pour entraîner un tel système de RDF. Notez que toutes les expérimentations sur la classification automatisée ont été conduites sur un sous échantillonnage des données : seuls les segments d'images identifiés au cours de la segmentation manuelle comme appartenant au *corail* ont été retenus. Ainsi, nous aurons une meilleure idée des performances du système de classification, sans considérer les erreurs de segmentation automatique qui pourraient influencer les résultats.

### 3.8.2.1 Évaluation de la capacité de classification en considérant deux métaclases

Lorsqu'un faible nombre d'échantillons est disponible pour chaque classe, il devient intéressant de faire un groupement parmi les classes et ainsi former une métaclasse. Comme discuté en 3.3.1.2 une distinction est souvent faite dans le monde de la biologie entre les coraux de type *acropora* et *non-acropora*. La présente expérimentation tentera de distinguer ces deux concepts en considérant deux groupements parmi les 31 classes échantillonnées. La classe *acropora* concerne 8 131 segments (27,1 %) et *non-acropora* 21 782 segments (72,9 %). Les résultats bruts de cette expérience sont présentés par le Tableau 3.5.

Tableau 3.5 Taux de classification bruts de classification (métaclases)

Itération	SVM	kPPV	Itération	SVM	kPPV
1	75,6 %	67,4 %	5	74,8 %	67,1 %
2	75,3 %	66,4 %	6	75,1 %	67,8 %
3	74,4 %	68,6 %	7	75,0 %	67,2 %
4	75,2 %	67,6 %	8	74,8 %	66,3 %

Pour cette expérience, le **taux de classification moyen** estimé est de **75,02 ± 0,30 %** pour le **SVM** et de **67,29 ± 0,62 %** pour le **kPPV**. Cette mesure de performance peut laisser croire que l'expérience est un succès. Toutefois, les matrices de confusion totales (somme des huit itérations) liées à la classification des spécimens parmi les classes *acropora* et *non-acropora* laissent entre-voir une autre conclusion (voir le Tableau 3.6).

En étudiant les matrices de confusion de chaque itération, on remarque de piètres performances pour la classe *acropora* qui admet une **précision** de **70,41 ± 1,64 %** et un **taux de rappel** de **14,13 ± 2,25 %**. Le classificateur commet donc beaucoup d'erreurs de classification au sujet de la classe *acropora*. Ce phénomène est fort probablement dû à la grande variabilité interclasses liée à l'amalgame de toutes les espèces du genre *acropora* dans une seule classe.

Tableau 3.6 Matrices de confusion de la classification par métaclasse

		Classe réelle (SVM)		Classe réelle (kPPV)	
		Acropora	Non-acropora	Acropora	Non-acropora
Classe estimée	Acropora	1 149	489	3 110	4 763
	Non-acropora	6 982	21 293	5 021	17 019

### 3.8.2.2 Évaluation de l'effet du nombre de classes sur les performances

L'approche par métaclasse étant peu concluante, nous avons considéré le problème à son niveau le plus fondamental : la classification des spécimens au niveau des espèces (le plus fin niveau taxinomique). Ainsi, la variabilité interclasses est à son minimum et est représentative de la variation naturelle expliquée par l'échantillonnage encyclopédique dont nous disposons.

Au cours de cette expérience, nous avons tenté d'évaluer la quantité d'espèces distinctes que nous pouvons arriver à départager en considérant un nombre fixe d'échantillons d'images. En débutant avec deux classes uniquement, jusqu'à six espèces distinctes. Enfin, pour chaque quantité de classes à départager, huit expériences ont été tentées, en modifiant les groupes d'espèces sélectionnées aléatoirement, afin de minimiser l'impact des particularités inhérentes à chaque comparaison. Pour cette expérience, toutes les données disponibles pour chaque classe ont été utilisées. Les taux de classification observés sont présentés graphiquement à la Figure 3.25. Sur cette figure, pour chaque nombre de classes considéré, les intervalles de confiance de Student sur la moyenne sont représentés par une barre d'erreur. Notez également qu'un segment de droite relie chaque point dans l'unique but de mieux visualiser la courbure de la relation qui unit le nombre de classes et le taux de classification. Les groupes de classes considérés pour chaque quantité d'espèces à départager sont détaillés en 3.3.1.3.



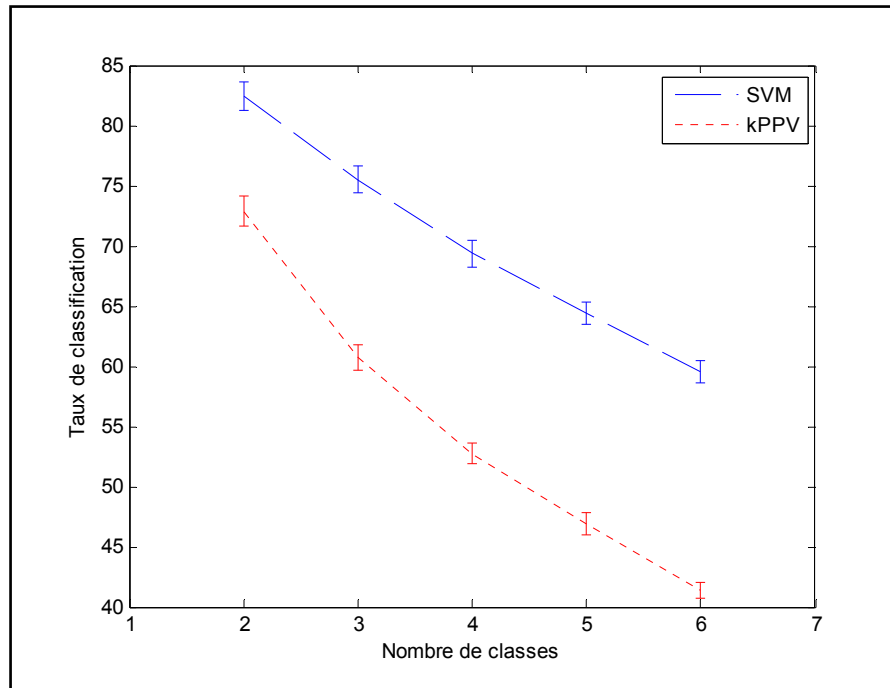


Figure 3.25 Taux de classification en fonction du nombre de classes.

Les petits intervalles de confiance observés sur les résultats démontrent que les classes choisies parmi nos données représentent une tâche de classification dont la difficulté reste similaire, même en considérant 8 échantillonnages de 2, 3, 4, 5 et 6 classes dont la sélection est purement aléatoire. Ce résultat est important pour nous, puisqu'il nous permet de conclure sur la difficulté du problème, indépendamment des classes considérées.

Le SVM admet un **taux de classification moyen** de  $82,42 \pm 1,19$  % lorsque deux classes sont considérées. Cette performance plus qu'acceptable dans le domaine du suivi des communautés benthiques laisse voir que le lot de descripteurs utilisé est suffisamment discriminant pour distinguer deux espèces de corail, peu importe la combinaison choisie dans notre échantillon. Toutefois, la décroissance rapide de la performance lorsque le nombre de classes considéré augmente confirme qu'il est impossible de bien discerner toutes les espèces du problème avec aussi peu de données par classes.

### 3.8.2.3 Évaluation de l'effet de la quantité de données par classe sur les performances

La quantité de données considérée est mise en évidence dans l'expérimentation précédente comme étant une variable importante du succès d'un système de classification. Afin de mieux visualiser la quantité d'information par classes qui serait nécessaire à entraîner un système pour la problématique complète, nous avons tenté de moduler cette variable dans le cadre de la prochaine expérience.

La quantité d'information utilisée peut se matérialiser par le nombre de segments d'image utilisés pour entraîner le classificateur. Toutefois, la nature variable de la taille des segments implique une incertitude quant à la stabilité statistique des descripteurs qui y sont extraits. Les descripteurs statistiques extraits d'un segment possédant une aire de 30 pixels carrés ne seront pas aussi fiables que ceux extraits d'un segment de l'image possédant une superficie de 3 000 pixels. Pour pondérer cet effet et pour utiliser une mesure qui est indépendante de l'algorithme de présegmentation non supervisée, nous avons choisi la surface utilisée du corail (en pixels) pour mesurer la quantité d'information utilisée.

Au cours de notre expérience, nous avons donc fait varier la surface utilisée pour chaque espèce de 100K pixels à 800K pixels. Cette méthode garantit une représentation similaire des espèces dans les bases de données d'entraînement, de validation et de test. De plus, pour compléter le portrait de la situation, nous avons conduit l'expérimentation en considérant de deux à six espèces distinctes. Toutefois, pour limiter le nombre de répliques nécessaires, les groupes d'espèces ayant obtenu les taux de classification médians au cours de l'expérimentation précédente ont été utilisés. Encore une fois, huit répliques ont été conduites, en variant aléatoirement chaque fois, les segments utilisés pour représenter chacune des espèces. La Figure 3.26 montre les taux de classification en fonction de la superficie d'image utilisée par espèce tels qu'obtenus pour le SVM et la Figure 3.27, pour le kPPV.

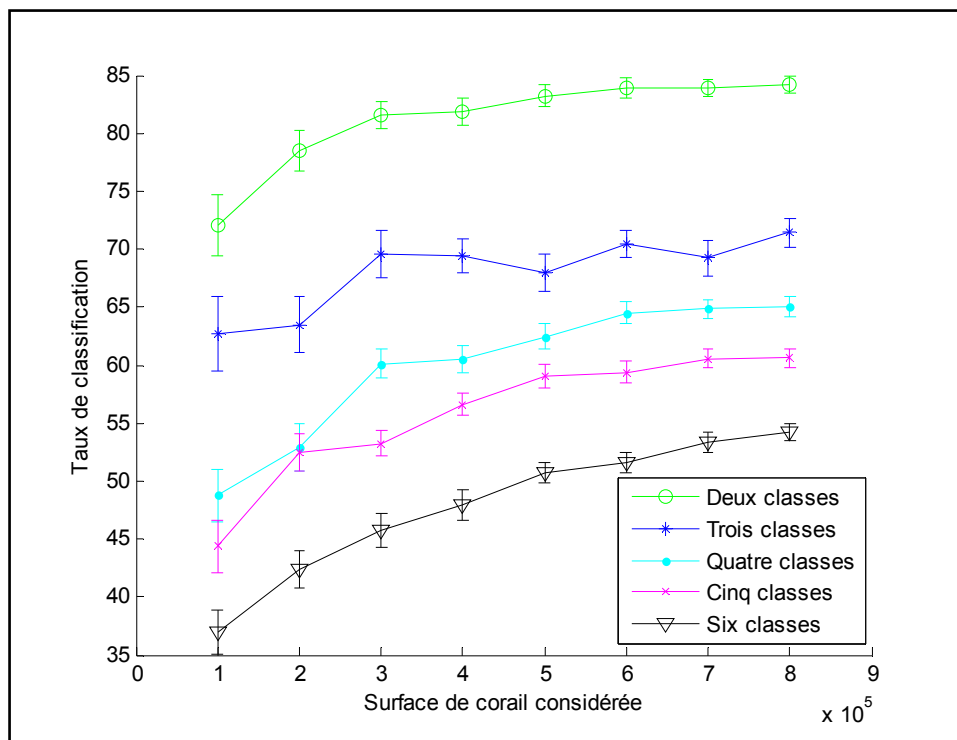


Figure 3.26 Taux de classification en fonction de la surface considérée pour le SVM.

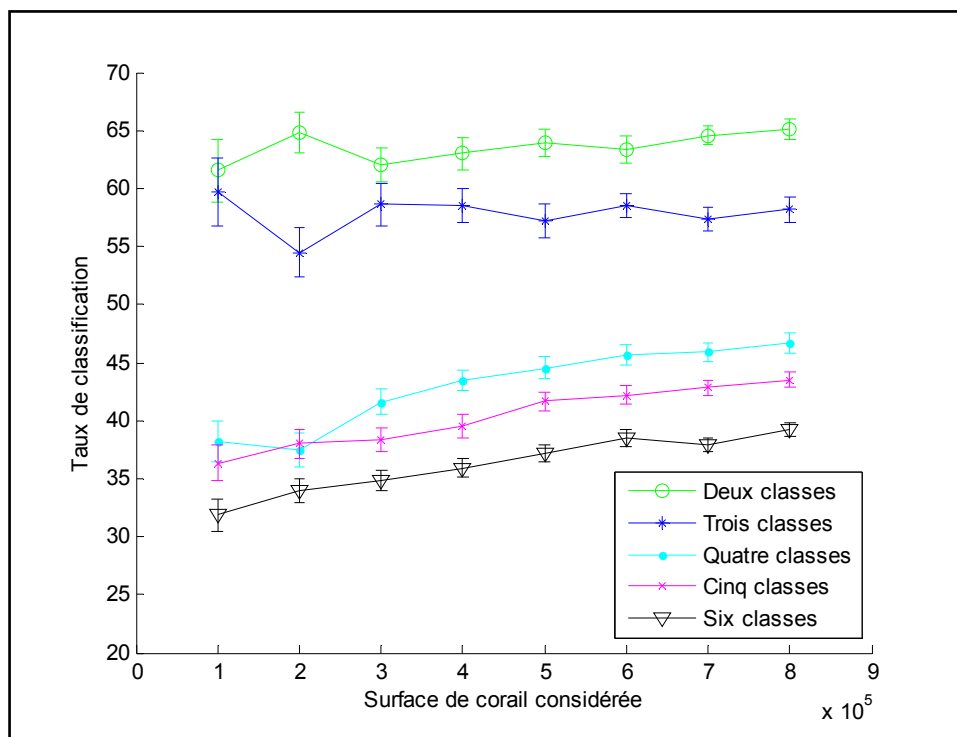


Figure 3.27 Taux de classification en fonction de la surface considéré pour le kPPV.

Les précédentes figures montrent que la relation qui unit la quantité d'information utilisée et le taux de classification semble tendre asymptotiquement vers un maximum. De plus, les performances observées à 800K pixels pour le SVM sont grossièrement comparables à celles découvertes en 3.8.2.2 où toutes les données disponibles sont utilisées (avec une superficie par espèce pouvant atteindre jusqu'à ~2 500K pixels, voir Figure 3.18). Pour bien saisir ce phénomène, il faut se rappeler que les segments d'images utilisés pour concevoir l'expérimentation proviennent de 6 à 8 spécimens distincts uniquement (voir 3.3). La variabilité naturelle de chaque espèce n'est donc pas totalement expliquée par l'échantillonnage. De ce fait, l'apprentissage est rapidement limité aux connaissances véhiculées par la banque d'images de *Corail ID* lorsque l'on accroît la quantité d'information considérée. Ce résultat implique que l'augmentation de la taille de la base de données pourrait s'avérer une solution au problème. Toutefois, ce n'est qu'une partie de la solution. Des médiums et des descripteurs plus discriminants pourraient également être nécessaires pour solutionner la problématique en réduisant la confusion qui règne entre les classes.

### **3.9 Discussion générale sur l'approche par description taxinomique complète**

L'expérimentation par description taxinomique complète nous a renseignés sur la faisabilité technologique d'une classification automatisée du corail à partir d'images numériques. La difficulté inhérente à la classification d'objets biologiques est visible dans nos résultats expérimentaux. Toutefois, les performances obtenues en 3.8.2.3 laissent croire que ce type de système pourrait voir le jour considérant qu'assez d'échantillons d'images sont acquis pour entraîner le système et qu'un lot de descripteurs discriminant y est extrait. L'atomicité des classes liée à la classification taxinomique est certes un avantage, ce qui valide notre première hypothèse de travail liée à cette portion de l'étude : *l'entraînement d'un système intelligent d'identification de corail est possible à l'aide d'une banque d'images classifiées à un niveau atomique.*

Par conséquent, *puisque les variabilités interclasses et intraclasse sont réduites par l'identification taxinomique, la quantité de données nécessaire est moins importante pour*

*assurer le bon fonctionnement du système de classification.* Cette hypothèse est également confirmée par les résultats de l'expérimentation 3.8.2.3 où un sous-échantillonnage des données disponibles à l'entraînement a été utilisé et où les performances tendent asymptotiquement vers celles obtenues considérant toutes les données disponibles en 3.8.2.2. De plus, on remarque que les performances sont grandement dégradées lorsque la variabilité intraclasse est augmentée par l'utilisation de métaclasse en 3.8.2.1.

Au niveau de la segmentation des spécimens, notre hypothèse stipulant que le sujet photographique de l'image peut être segmenté correctement, *étant donné l'information véhiculée par le point focal de l'image et l'intention du photographe*, a aussi été validée. La pertinence des résultats de segmentation obtenus quantitativement et qualitativement démontre bien que ce type d'opération est automatisable, même en considérant la grande complexité de certaines scènes sous-marines.

Malgré la validation des hypothèses par nos expériences sur la banque d'images de *Coral ID*, les performances obtenues restent sous-optimales. Plusieurs problèmes pourraient donc être résolus pour réduire davantage l'erreur de classification. En premier lieu, les images acquises en contexte encyclopédique ne sont pas idéales à l'entraînement d'un tel système. Il serait préférable d'utiliser un protocole de photographie qui uniformise l'échelle et l'angle de prise de vue et admet une résolution d'images supérieure à celle de *Coral ID*. De plus, la quantité d'images par espèce de corail à départager devrait être augmentée significativement pour bien saisir toute la variabilité inhérente à chaque espèce de corail. Enfin, l'information spectrale et géométrique pourrait être captée sur les sites d'études. L'utilisation des technologies d'imagerie multispectrale et de vision stéréoscopique serait un atout pour mieux discriminer les nombreuses espèces d'intérêt. Une étude plus poussée sur le sujet devra toutefois être conduite pour déterminer dans quelles conditions ces technologies complexes peuvent bénéficier à la problématique.

Suite à ces considérations sur les données, d'autres recommandations peuvent être émises au sujet du système de classification lui-même. Dans un premier temps, une étude plus poussée

au sujet des descripteurs de morphologie et de texture serait bénéfique. Selon nos observations, la morphologie des spécimens est une information très corrélée avec l'espèce. Dans ce contexte, l'amélioration de l'algorithme de présegmentation non supervisée et l'ajout de descripteurs morphologiques serait à considérer. Enfin, pour caractériser la morphologie des coraux à plus petite échelle, les descripteurs de textures restent un sujet d'intérêt. La topologie des polypes, partie du corail de petite taille sur nos images, est l'information la plus discriminante ciblée par *Coral ID* (Veron et Stafford-Smith, 2002) et peut certainement être étudiée du point de vue de la texture. Au final, l'implémentation d'un mécanisme de rejet, tel que présenté en 2.5.3, serait envisageable pour compléter manuellement la classification des spécimens qui présentent une difficulté accrue.

En somme, les résultats obtenus et les méthodes utilisées au cours des précédentes expérimentations sont comparables à la littérature (voir section 1.1). Nous avons démontré que le problème de classification du corail à un niveau d'identification taxinomique est solvable en partie par l'implémentation d'un mécanisme de présegmentation non supervisée des images et l'entraînement de classificateurs pour traiter les segments d'images extraits. Pour porter ces conclusions à l'ensemble de la problématique et ainsi considérer globalement les 286 espèces de corail envisagées, un travail de collecte massive de données et d'extraction d'information discriminante devra être complété.

## CONCLUSION

Dans ce mémoire, nous avons implémenté l'automatisation de la classification de spécimens de corail à partir d'images numériques en utilisant des méthodes issues de la théorie sur la vision et l'intelligence artificielle. L'automatisation de cette tâche a été considérée dans le but d'alléger le travail des biologistes qui œuvrent actuellement à l'étude de la biodiversité et de son évolution relative aux perturbations de l'homme sur l'environnement.

Au cours des travaux conduits sur cette problématique, deux approches distinctes d'identification des coraux ont été mises à l'épreuve. Dans un premier temps, la description de l'aspect visuel des coraux à l'aide de la terminologie des formes benthiques a été tentée. Au cours de cette partie de l'étude, divers descripteurs de texture et de couleur ont été utilisés pour implémenter l'identification des formes benthiques à partir d'images numériques non segmentées. L'automatisation de la reconnaissance de ces formes benthiques sur les images numériques s'est avérée infructueuse. Les deux principales raisons qui expliquent les résultats obtenus sont la grande *variabilité intraclasse* implicite à la classification par formes benthiques et l'absence d'*information morphologique* parmi les descripteurs utilisés, dans un contexte d'identification de formes morphologiques benthiques. Aucune information morphologique n'avait été extraite des images puisqu'elles représentent un défi de segmentation sans précédent, dû à la nature très variée des objets et de leur environnement. Par contre, l'outil de segmentation semi-automatisée que nous avons développé par la suite pourrait aider la segmentation des entités dans les images et ainsi favoriser la mise en place d'une base de données où les objets sont segmentés.

Pour faire suite à la problématique sur les formes benthiques, l'identification au niveau taxinomique des coraux a été tentée à partir d'images numériques segmentées automatiquement et divers descripteurs de texture, de couleur et de morphologie. Cette description taxinomique, pourtant plus difficile à accomplir pour l'homme que celle par formes benthiques, s'est avérée plus facile à traiter par vision et intelligence artificielle. L'*atomicité* des classes taxinomiques et la *segmentation* des images sont les deux principales

contributions au succès de cette approche. Il est donc important d'utiliser le niveau de classification qui possède la plus fine granularité possible. À cet égard, le regroupement de classes au sein d'une métaclasse n'est pas une stratégie supportée par nos résultats expérimentaux. En plus de traiter du problème de la classification, la segmentation automatisée des images de corail a été implémentée au cours de l'étude sur la description taxinomique. À l'aide des outils développés, un algorithme de segmentation fonctionnel a pu être implémenté. Avec cet outil, les coraux peuvent facilement être repérés automatiquement dans un environnement complexe.

En conclusion de l'étude sur la description taxinomique complète, l'hypothèse de travail générale est supportée par les résultats: *l'identification visuelle à l'échelle macroscopique est possible pour départager la majorité des spécimens de corail*. Cette hypothèse a toutefois été soumise à une condition importante au cours des travaux sur les formes benthiques, qui reste toujours souhaitable, même avec la représentation taxinomique : *assez de données doivent être disponibles pour caractériser la relation entre le stimulus visuel et l'identité du spécimen*.

Les résultats suggèrent également qu'il est possible, dans les conditions énumérées ci-haut, de relever le défi lié au suivi écologique automatisé tout en minimisant la complexité technologique. L'utilisation de simples caméras numériques en contexte de plongée pourrait être suffisante pour effectuer un suivi des communautés benthiques. Toutefois, l'ajout d'informations supplémentaires, telles la *topologie tridimensionnelle* et la *réponse spectrale* pourrait s'avérer un atout majeur pour repousser les limites du système et ainsi départager un grand nombre de formes benthiques ou d'espèces avec une précision acceptable, tel qu'exploré par certains travaux de la littérature.

De plus, il serait préférable de considérer indépendamment la problématique de segmentation et de classification des spécimens. Un lot de descripteurs plus adapté à chacune des tâches serait un atout et maximiserait la performance liée à ces deux étapes importantes du processus. En effet, ces deux problématiques sont assez vastes pour faire l'objet d'une



recherche indépendante. Plus encore, il serait bénéfique d'intégrer les spécifications d'ingénierie du système de reconnaissance de formes au protocole d'acquisition des images pour s'assurer, en premier lieu, d'une stabilité de l'échelle et de l'angle des prises de vue.

Ce projet de maîtrise se veut une première itération dans le processus de développement d'une solution viable. Ce n'est qu'une preuve de concept, une recherche exploratoire. Nous sommes conscients que cette problématique est très vaste, qu'une petite partie d'elle, tel par exemple la présegmentation non supervisée d'une image, peut elle-même faire l'objet d'une étude de la même envergure. C'est pourquoi nous n'avons pas poussé plus loin les recherches pour estimer la performance espérée pour la problématique complète qui considère la segmentation automatisée suivie de la classification automatisée.

En somme, force est de constater que l'automatisation du suivi de l'évolution des communautés benthiques est chose possible. Les résultats obtenus suite à l'étude laissent voir qu'avec un échantillonnage adéquat et avec l'acquisition d'informations discriminantes, un problème aussi complexe que l'identification visuelle de l'espèce d'un spécimen corallien pourrait être accompli. Plus encore, l'architecture logicielle développée au cours de ce travail pourrait être réutilisée et personnalisée pour répondre à d'autres besoins. Par exemple, l'analyse de la couverture de forêts par imagerie aérienne est une des nombreuses applications possibles à un tel système de segmentation et classification de régions par vision artificielle et reconnaissance de formes.



## RECOMMANDATIONS

Au cours de l'élaboration de ce vaste projet d'automatisation, plusieurs décisions ont été prises pour limiter la recherche à un cadre théorique bien défini. Plusieurs pistes de solution n'ont pu être explorées faute de temps ou de données. Cette dernière section du mémoire présentera, pour chacune des six étapes du processus d'identification automatisé par vision et intelligence artificielle (voir Figure 1.1), diverses recommandations ou pistes de recherches jugées pertinentes.

Dans un premier temps, au sujet de l'acquisition des informations brutes, plusieurs améliorations pourraient être considérées. L'acquisition d'informations supplémentaires sur le terrain pourrait s'avérer une étape importante pour implémenter un système d'identification de coraux à grande échelle, qui considère toutes les classes nécessaires à l'échantillonnage. L'information tridimensionnelle de la scène pourrait être acquise par sonar, par balayage infrarouge ou laser ou encore par stéréovision. Cette information sur la topologie du substrat marin et des organismes vivants qu'on y trouve est sans aucun doute très discriminante quant à la nature des objets à identifier. L'information tridimensionnelle aiderait à compléter le portrait morphologique d'un spécimen de corail. Cette morphologie, selon les enseignements de *Coral ID* (Veron et Stafford-Smith, 2002) est très fortement liée à l'espèce du spécimen en question. Enfin, dans l'optique d'une classification par formes benthiques, la morphologie tridimensionnelle du spécimen et l'information spectrale pourraient également être très utiles à l'identification des spécimens. Plus encore, la segmentation à l'aide des bandes spectrales adéquates pourrait être grandement simplifiée, comme le proposent plusieurs écrits de la littérature présentés au premier chapitre.

D'autres recommandations plus générales au sujet de l'acquisition des données ont également été mises en évidence au cours de l'étude de cette problématique. Le protocole de prise de vue sur le terrain doit prendre en compte les requis d'ingénierie liés à l'implémentation de la reconnaissance automatisée des spécimens coralliens. L'angle et la distance de prise de vue doivent être normalisés pour maximiser les performances. De plus,

la variation dans l'éclairage devrait également être limitée au maximum. L'emploi d'un éclairage d'appoint submersible pourrait s'avérer une solution au problème et ainsi limiter l'impact de la météo sur la constance de l'illumination des spécimens. Plus encore, la date et l'heure des prises de vues sur le terrain devraient être normalisées pour éviter la variabilité liée à l'extension de tissus de coraux (cycle jour / nuit) ou encore liée à la période de reproduction de ces êtres vivants. Ces variations temporelles du corail doivent être prises en compte, car elles influencent énormément l'apparence de certaines espèces. Enfin, pour ajouter plus d'information discriminante, l'information géographique (la position GPS par exemple) et l'information bathymétrique (profondeur et position sur le littoral de l'observation) pourraient être ajoutées aux données recueillies pour chaque image.

Enfin, la quantité d'échantillons acquis doit être de beaucoup supérieure à celle observée dans les deux bases de données utilisées dans ce travail. L'échantillonnage doit couvrir suffisamment de spécimens pour permettre une modélisation de la variabilité de chaque espèce. Dans un contexte idéal, l'échantillonnage devrait prendre en compte la variabilité de forme pour une même espèce, telle qu'imposée par le milieu de croissance et la variabilité de couleur relative à la symbiose avec différentes espèces d'algues cellulaires. De plus, la représentation choisie pour classifier les spécimens devrait favoriser l'atomicité des classes. L'atomicité des classes minimise la variation intraclasse et par le fait même, la confusion entre les classes dans l'espace de description. L'approche taxinomique est donc favorable.

Au niveau de l'échantillonnage pour l'entraînement d'un algorithme de RDF, il semble avantageux de favoriser la répartition naturelle des classes. De plus, nos observations démontrent que la validation croisée est un excellent moyen pour structurer les expérimentations (lorsque le but est d'évaluer la performance du système).

Au sujet de la segmentation, diverses méthodes avancées pourraient être étudiées et appliquées au problème des coraux. Le corail est une colonie d'individus nommés *polypes* qui s'étendent peu à peu sur le substrat benthique. De ce point de vue, plusieurs coraux forment peu à peu une texture où divers motifs (les polypes) se répètent dans l'espace.

Certaines méthodes de la littérature permettent la segmentation du « *texel* » de la texture, elles retrouvent et reconstruisent l'*élément de texture* qui se répète dans une image (Todorovic et Ahuja, 2009). Ce type d'algorithme permettrait donc d'isoler le polype moyen du corail et d'y extraire certains descripteurs. Il serait intéressant de tester cette idée dans une version future du système de reconnaissance puisque toujours selon *Coral ID*, le polype est un élément très discriminant de l'espèce. Il serait également intéressant d'obtenir une arborescence de segmentation dans laquelle l'imbrication des diverses composantes de l'image est consignée. Ce genre de segmentation permettrait ensuite d'extraire une description morphologique hiérarchique du spécimen, ce qui pourrait améliorer la capacité de discriminer entre les espèces. D'autres recommandations générales concernent la présegmentation des images. Cet algorithme pourrait être amélioré de façon à réduire le nombre de segments résultants de son exécution et ainsi éviter d'obtenir de très petites zones où l'extraction de certains descripteurs devient instable. Enfin, comme mentionné précédemment, il serait intéressant d'étudier quelles bandes spectrales peuvent être utiles à la segmentation des coraux. En quelques opérations de seuillage, il pourrait être très évident ainsi de départager le corail d'un arrière-plan sédimentaire tel le sable et les rochers.

Un travail plus poussé sur l'extraction de caractéristiques est une des recommandations que nous jugeons la plus importante. Pour permettre la discrimination des nombreuses espèces de corail, l'étude de caractéristiques liées à la morphologie du corail et la géométrie des polypes devrait être la prochaine action à entreprendre pour continuer ce projet. À ce sujet, il serait intéressant aussi d'optimiser l'ensemble de caractéristiques utilisé indépendamment pour le problème de la segmentation et de la classification.

Finalement, l'optimisation, l'entraînement et la généralisation, soit les dernières étapes du processus pourront également faire l'objet d'améliorations. Toutefois, nous considérons le SVM à noyau RBF optimisé comme un excellent choix puisqu'il est rapide à entraîner, il possède une bonne capacité de généralisation, est performant et est simple à paramétrer. Il serait intéressant par contre de tenter l'optimisation des métaparamètres du SVM à partir d'une autre métrique que le taux de classification. L'aire sous la courbe ROC (*receiver*

*operator characteristic*) ou les métriques de précision et rappel pourraient favoriser d'autres paramétrisations optimales qui possèdent des caractéristiques intéressantes en généralisation. Au niveau de l'architecture du système de classification également, des innovations pourraient être apportées. Pour spécialiser davantage les classificateurs, un système de détecteurs, un pour chacune des espèces considérées, pourrait être entraîné avec un lot spécialisé de caractéristiques. Enfin comme dernière recommandation au sujet de l'architecture, un mécanisme de rejet pourrait être ajouté en sortie. Ainsi lorsque le système n'est pas confiant sur la décision au sujet d'un certain spécimen, la classification manuelle de ce spécimen pourrait être accomplie. De cette façon, même avec un faible nombre d'échantillons utilisés pour entraîner le système, une bonne partie du travail pourrait être effectuée automatiquement.

Voilà ce qui conclut les différentes recommandations pour les travaux futurs sur ce système. En somme, les deux points les plus critiques à améliorer sont en premier lieu la quantité et la qualité des images de la base de données d'entraînement et enfin l'extraction de descripteurs morphologiques pertinents à la classification des coraux, inspirés si possible des travaux en biologie sur le sujet, le tout en maintenant l'atomicité des classes considérées.

## LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES

- Abramoff, M.D., P.J. Magelhaes et S.J. Ram. 2004. « Image Porcessing with ImageJ ». *Biophotonics International*, vol. 11, n° 7, p. 36-42.
- Ahuja, N., et S. Todorovic. 2007. « Extracting Texels in 2.1D Natural Textures ». In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*. p. 1-8.
- Bellman, Richard Ernest. 1961. *Adaptive Control Processes : A Guided Tour*. Princeton (NJ): Princeton University Press, 255 p.
- Bouchard, Jonathan. 2008a. *Projet de fin d'étude : Optimisation des méthodes de classification*. Exigence du cours GPA792. Montréal: École de technologie supérieure, 22 p. < <http://profs.etsmtl.ca/jalandry/Recherche/Bouchard/> >.
- Bouchard, Jonathan. 2008b. *Rapport technique : Préparation des données pour la classification d'images de coraux*. Exigence du cours SYS821. Montréal: École de technologie supérieure, 20 p. < <http://profs.etsmtl.ca/jalandry/Recherche/Bouchard/> >.
- Bouchard, Jonathan. 2009a. *Lecture dirigée : Méthodes d'extraction de caractéristiques et de segmentation non supervisée pour textures naturelles*. Exigence du cours MTR871. Montréal: École de technologie supérieure, 27 p. < <http://profs.etsmtl.ca/jalandry/Recherche/Bouchard/> >.
- Bouchard, Jonathan. 2009b. *Rapport technique : Implémentation d'un système de détection et de classification de coraux*. Exigence du cours SYS863. Montréal: École de technologie supérieure. < <http://profs.etsmtl.ca/jalandry/Recherche/Bouchard/> >.
- Bouchard, Jonathan. 2009c. *Rapport technique : Implementation d'un systeme multi-étages pour la classification de coraux*. Exigence du cours SYS843. Montréal: École de technologie supérieure, 26 p. < <http://profs.etsmtl.ca/jalandry/Recherche/Bouchard/> >.
- Bouchard, Jonathan. 2009d. *Revue de littérature : La fusion d'information en reconnaissance biométrique*. Exigence du cours SYS863. Montréal: École de technologie supérieure, 26 p. < <http://profs.etsmtl.ca/jalandry/Recherche/Bouchard/> >.
- Bourque, Pierre-André. 2009. « Planète Terre : Les grands cycles biogéochimiques et les changements climatiques », Support de cours. En ligne et télédiffusé sur Canal Savoir. Québec (Qué.): Le département de Géologie et de Génie géologique de l'Université Laval. < [http://www2.ggl.ulaval.ca/personnel/bourque/intro.pt/planete\\_terre.html](http://www2.ggl.ulaval.ca/personnel/bourque/intro.pt/planete_terre.html) >.
- Boyce, Daniel G., Marlon R. Lewis et Boris Worm. 2010. « Global phytoplankton decline over the past century ». *Nature*, vol. 466, n° 7306, p. 591-596.

- Bradski, G. 2000. *The OpenCV library*. Logiciel. Bellevue (WA): Dr. Dobb's Journal of Software Tools. < <http://opencv.willowgarage.com> >.
- Chang, Chih-Chung, et Chih-Jen Lin. 2001. *LIBSVM : A library for support vector machines*. < <http://www.csie.ntu.edu.tw/~cjlin/libsvm> >.
- Chih-Wei, Hsu, et Lin Chih-Jen. 2002. « A comparison of methods for multiclass support vector machines ». *Neural Networks, IEEE Transactions on*, vol. 13, n° 2, p. 415-425.
- Chow, C. 1970. « On optimum recognition error and reject tradeoff ». *Information Theory, IEEE Transactions on*, vol. 16, n° 1, p. 41-46.
- Cleto L., Nañola, Jr., Angel C. Alcalá, Porfirio M. Aliño, Hazel O. Arceo, Wilfredo L. Campos, Edgardo D. Gomez, Wilfredo Y. Licuanan, Miledel C. Quibilan, Andre J. Uychiaoco et Alan T. White. 2004. « Philippines ». In *Status of Coral Reefs in East Asian Seas Region : 2004*, sous la dir. de Center, Japan Wildlife Research. p. 43-51. Shitaya (Tokyo).
- Cooley, JW, et JW Tukey. 1965. « An algorithm for the machine calculation of complex Fourier series ». *Mathematics of computation*, vol. 19, n° 90, p. 297-301.
- Cortes, Corinna, et Vladimir Vapnik. 1995. « Support-vector networks ». *Machine Learning*, vol. 20, n° 3, p. 273-297.
- Darwin, Charles. 1837. *Notebook B : Transmutation of species*. En ligne. < <http://darwin-online.org.uk> >.
- Diaz, JA, et RE Torres. 2007. « Classification of Underwater Color Images with Applications in the Study of Deep Coral Reefs ». In. Vol. 2, p. 610-613. IEEE.
- Duda, Richard O., Peter E. Hart et David G. Stork. 2001. *Pattern Classification*, 2e éd. New York (NY): John Wiley & Sons Inc., 654 p.
- Eberhart, R. C., et J. Kennedy. 1995. « A new optimizer using particle swarm theory ». In *Proceedings of the Sixth International Symposium on Micromachine and Human Science* (Nagoya, Japon, 4-6 octobre 1995). p. 39-43.
- English, Susan A., V. J. Baker, Clive R. Wilkinson, Australian Institute of Marine Science et ASEAN-Australia Marine Science Project. 1994. *Survey manual for tropical marine resources*. Townsville (Australie): Australian Institute of Marine Science, 368 p.
- Forsman, Zac, Daniel Barshis, Cynthia Hunter et Robert Toonen. 2009. « Shape-shifting corals : Molecular markers show morphology is evolutionarily plastic in *Porites* ». *BMC Evolutionary Biology*, vol. 9, n° 1, p. 45.



- Fumera, Giorgio, Fabio Roli et Giorgio Giacinto. 2000. « Reject option with multiple thresholds ». *Pattern Recognition*, vol. 33, p. 2099-2101.
- Gleason, ACR, RP Reid et KJ Voss. 2007. « Automated classification of underwater multispectral imagery for coral reef monitoring ». In *Oceans 2007* (Vancouver, 29 sept. - 4 oct. 2007). p. 1-8. Citeseer.
- Gonzalez, Rafael C., et Richard E. Woods. 2008. *Digital Image Processing*, 3e éd. Upper Saddle River (New Jersey): Pearson Prentice Hall, 954 p.
- Guay, Sébastien. 2007. *Projet de fin d'étude : Optimisation de la taille des régions d'intérêt dans des images non segmentées pour améliorer la performance de classificateurs appliqués aux objets biologiques*. Exigence du cours GPA792. Montréal: École de technologie supérieure, 24 p.
- Hall, Mark, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann et Ian H. Witten. 2009. « The WEKA Data Mining Software : An Update ». *SIGKDD Explorations*, vol. 11, n° 1, p. 10-18.
- Haralick, Robert M., K. Shanmugam et Its'Hak Dinstein. 1973. « Textural Features for Image Classification ». *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 3, n° 6, p. 610-621.
- Hodgson, G., W. Kiene, J. Mihaly, J. Liebler, C. Shuman et L. Maun. 2004. *Reef Check Instruction Manual: A Guide to Reef Check Coral Reef Monitoring*. Reef Check, Institute of the Environment, University of California at Los Angeles, 86 p. < <http://www.reefcheck.org> >.
- Holland, H. D. 2006. « The oxygenation of the atmosphere and oceans ». *Philosophical Transaction of the Royal Society B : Biological Science*, vol. 361, n° 1470 (29 juin), p. 903-15.
- Holland, John H. 1992. *Adaptation in natural and artificial systems : An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. Cambridge (Massachusetts): MIT press Cambridge.
- Hsu, Chih-Wei, Chih-Chung Chang et Chih-Jen Lin. 2003. « A Practical Guide to Support Vector Classification ». Taipei 106, Taiwan: < <http://www.csie.ntu.edu.tw/~cjlin> >.
- Johnson-Roberson, M, S Kumar et S Willams. 2007. « Segmentation and classification of coral for oceanographic surveys: a semi-supervised machine learning approach ». In., p. 1-6. IEEE.

- Johnson-Roberson, M., S. Kumar, O. Pizarro et S. Willams. 2006. « Stereoscopic imaging for coral segmentation and classification ». In *OCEANS 2006*. p. 1-6.
- Journaux, Ludovic, Marie-France Destain, Johel Miteran, Alexis Piron et Frederic Cointault. 2008. « Texture classification with generalized fourier descriptors in dimensionality reduction context: An overview exploration ». In. Vol. 5064 LNAI, p. 280-291. Coll. « Lecture Notes in Computer Science ». Paris, France: Springer Verlag. < [http://dx.doi.org/10.1007/978-3-540-69939-2\\_27](http://dx.doi.org/10.1007/978-3-540-69939-2_27) >.
- Landgrebe, Thomas, Pacl, Pavel k, David M. J. Tax et Robert P. W. Duin. 2005. *Optimising two-stage recognition systems*. 3541. Seaside, CA., United states: Springer Verlag, 206-215 p.
- Landry, Jacques-André, et Jonathan Bouchard. 2007. « Énoncés de laboratoire du cours de vision artificielle ». Montréal (QC): École de technologie supérieure. < <https://cours.etsmtl.ca/gpa659/> >.
- Lepage, Richard, et Basel Solaiman. 2003. *Les réseaux de neurones artificiels et leurs applications en imagerie et en vision par ordinateur*. Montréal (QC): École de technologie supérieure, 446 p.
- Lesser, MP, et CD Mobley. 2007. « Bathymetry, water optical properties, and benthic classification of coral reefs using hyperspectral remote sensing imagery ». *Coral Reefs*, vol. 26, n° 4, p. 819-829.
- Levasseur, Yan. 2008. « Techniques de l'intelligence artificielle pour la classification d'objets biologiques dans des images bidimensionnelles ». Montréal, École de technologie supérieure, 172 p.
- Licuanan, Wilfredo Y. 2001. *Benthic life-forms : Relevant to the Philippine Pacific coast*. Adaptée de Page et al. 2001 AIMS SOP #2 rev. 2. Manilla (Philippines)Présentation Microsoft Powerpoint, 41 diapositives.
- Licuanan, Wilfredo Y. 2008. « Taxas of Philippines ». Électronique (Microsoft Excel). Université De La Salle, Manille (Philippines)
- Liu, YC, WH Chan et YQ Chen. 2002. « Automatic white balance for digital still camera ». *Consumer Electronics, IEEE Transactions on*, vol. 41, n° 3, p. 460-466.
- MacQueen, J. 1967. « Some Methods for classification and Analysis of Multivariate Observations ». In *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*. p. 281–297. 1. Los Angeles (CA): University of California Press.

- Marcos, M. 2005. « Classification of coral reef images from underwater video using neural networks ». *Optics Express*, vol. 13, n° 22, p. 8766-8771.
- Maritorena, S, A Morel et B Gentili. 1994. « Diffuse reflectance of oceanic shallow waters: influence of water depth and bottom albedo ». *Limnology and Oceanography*, vol. 39, n° 7, p. 1689-1703.
- McLaren, K. 1976. « XIII—The Development of the CIE 1976 (L\* a\* b\*) Uniform Colour Space and Colour-difference Formula ». *Journal of the Society of Dyers and Colourists*, vol. 92, n° 9, p. 338-341.
- Mehta, A, E Ribeiro, J Gilner et R van Woesik. 2007. « Coral reef texture classification using support vector machines ». In.
- Meyer, Fernand. 1994. « Topographic distance and watershed lines ». *Signal Process.*, vol. 38, n° 1, p. 113-125.
- Milgram, Jonathan. 2007. « Contribution à l'intégration des machines à vecteurs de support au sein des systèmes de reconnaissance de formes : Application à la lecture automatique de l'écriture manuscrite ». Montréal (QC), École de technologie supérieure, 155 p.
- Miller, David. 2007. « Corals : More complex than you? ». En ligne. Townsville, Australie: Science Daily. < [http://www.coralcoe.org.au/news\\_stories/coralgenes.html](http://www.coralcoe.org.au/news_stories/coralgenes.html) >.
- Office québécois de la langue française. 2010. *Le grand dictionnaire terminologique*. En ligne. < <http://www.granddictionnaire.com> >.
- Ojala, Timo, Matti Pietikäinen et David Harwood. 1996. « A comparative study of texture measures with classification based on featured distributions ». *Pattern Recognition*, vol. 29, n° 1, p. 51-59.
- ONU. 19 janvier 2007. *Résolution adoptée par l'Assemblée générale : 2010, Année internationale de la biodiversité*. 2 p. < <http://www.unesco.ch/fileadmin/documents/pdf/resolutions/ares61203fr.pdf> >.
- Otsu, Nobuyuki. 1979. « A Threshold Selection Method from Gray-Level Histograms ». *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, n° 1, p. 62-66.
- Page, C., G. Coleman, R. Ninio et K. Osborne. 2001. *Standard Operational Procedure Number 7 : Surveys of benthic reef communities using underwater video*. Townsville (Australie): Australian Institute of Marine Science, 45 p. < <http://www.aims.gov.au> >.

- Palomar, Mike Joshua S., et Eduardo Mendoza. 2006. *Comparison of image analysis techniques to describe coral reef cover at a Philippine site*. Quezon City (Philippines): University of the Philippines.
- Quinlan, J. Ross. 1993. *C4.5 : Programs for Machine Learning*. Maurgan Kaufmann Publishers Inc., 302 p.
- Radio-Canada. 2010. « Déclin marqué du plancton ». En ligne. < <http://www.radio-canada.ca/nouvelles/environnement/2010/07/28/002-phytoplancton-ocean.shtml> >. Consulté le 29 juillet 2010.
- Rivera Maldonado, Francisco José. 2005. « Segmentation of underwater multispectral images with applications in the study of coral reefs ». In. Vol. 43, p. 284. ProQuest Information & Learning, 300 N Zeeb Rd. PO Box 1346 Ann Arbor MI 48106-1346 USA.
- Roach, John. 2004. « Source of Half Earth's Oxygen Gets Little Credit ». En ligne. National Geographic News < [http://news.nationalgeographic.com/news/2004/06/0607\\_040607\\_phytoplankton.html](http://news.nationalgeographic.com/news/2004/06/0607_040607_phytoplankton.html) >.
- Rockefeller University. 2003. « Deadly coral toxin exposes ion pump's deepest secret ». *news & notes* (New-York). En ligne. 17 janvier. < [www.rockefeller.edu/pubinfo/news\\_notes/nn\\_011703.pdf](http://www.rockefeller.edu/pubinfo/news_notes/nn_011703.pdf) >.
- Ross, Arun, et Anil Jain. 2003. « Information fusion in biometrics ». *Pattern Recognition Letters*, vol. 24, n° 13, p. 2115-2125.
- Sheppard, Charles. 2007. *CoralPedia : Your guide to Caribbean corals and sponges*. En ligne. < <http://coralpedia.bio.warwick.ac.uk> >. Consulté le 18 octobre 2010.
- Soriano, M., S. Marcos, C. Saloma, M. Quibilan et P. Alino. 2001. « Image classification of coral reef components from underwater color video ». In *OCEANS, 2001. MTS/IEEE Conference and Exhibition*. Vol. 2, p. 1008-1013 vol.2.
- Still, Michael. 2006. *The Definitive Guide to ImageMagick*. Coll. « The Expert's Voice ». New-York (NY): Springer-Verlag, 335 p.
- Student. 1908. « The Probable Error of a Mean ». *Biometrika*, vol. 6, n° 1 (Mars), p. 1-25.
- Todorovic, Sinisa, et Narendra Ahuja. 2009. « Texel-based Texture Segmentation ». In *IEEE International Conference on Computer Vision*. Kyoto, Japan.
- Vanhamel, I, I Pratikakis et H Sahli. 2003. « Multiscale gradient watersheds of color images ». *Image Processing, IEEE Transactions on*, vol. 12, n° 6, p. 617-626.

- Veron, John Edward Norwood. 2000. *Corals of the world*, 3 vol. Townsville, Australie: Australian Institute of Marine Science, 1382 p.
- Veron, John Edward Norwood, et Marie G. Stafford-Smith. 2002. *Coral ID : An electronic key to the zooxanthellate scleractinian corals of the world*, version. 1.1. CD-ROM. Townsville, Australie: Australian Institute of Marine Science.
- Voronoi, G. 1908. « Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Premier mémoire. Sur quelques propriétés des formes quadratiques positives parfaites ». *Journal für die reine und angewandte Mathematik (Crelle's Journal)*, vol. 1908, n° 133, p. 97-102.
- Wikimedia Foundation. 2010. *Wikipedia : l'encyclopédie libre*. En ligne. < [www.wikipedia.org](http://www.wikipedia.org) >. Consulté le juillet 2010.
- Wilkinson, Clive. 2008. *Status of Coral Reefs of the World : 2008*. Townsville (Australie): Global Coral Reef Monitoring Network et Reef and Rainforest Research Center, 296 p.
- Witten, Ian H., et Eibe Frank. 2005a. *Data Mining : Practical Machine Learning Tools and Techniques*, Second Edition. San Francisco (CA): Elsevier, 525 p.
- Witten, Ian H., et Eibe Frank. 2005b. *Data Mining: Practical machine learning tools and techniques*, 2. San Francisco: Morgan Kaufmann, 525 p.