



HAL
open science

Segmentation d'images par morphologie mathématique et classification de données par réseaux de neurones : Application à la classification de cellules en cytologie des séreuses

Olivier Lezoray

► **To cite this version:**

Olivier Lezoray. Segmentation d'images par morphologie mathématique et classification de données par réseaux de neurones : Application à la classification de cellules en cytologie des séreuses. Traitement des images [eess.IV]. Université de Caen Basse-Normandie, 2000. Français. NNT: . tel-01100427

HAL Id: tel-01100427

<https://hal.science/tel-01100427>

Submitted on 6 Jan 2015

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

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



Numéro attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--

UNIVERSITE de CAEN/BASSE-NORMANDIE

U.F.R de Sciences

ECOLE DOCTORALE Structure, Information, Matière et Matériaux

THESE

Présentée par

M. LEZORAY Olivier

Et soutenue

Le 13/01/2000

En vue de l'obtention du

DOCTORAT de L'UNIVERSITE de CAEN

Spécialité : Informatique

(Arrêté du 30 mars 1992)

**Segmentation d'images couleur par morphologie
mathématique et classification de données par réseaux de
neurones : application à la classification de cellules en
cytologie des séreuses**

devant le jury composé de

M. Pierre Bonton	Rapporteur	Professeur
M. Hubert Cardot	Examineur	Maître de Conférences
M. Hubert Elie	Examineur	Médecin, Praticien Hospitalier
M. Abderrahim Elmoataz	Examineur	Maître de Conférences
M. Alain Faure	Rapporteur	Professeur
M ^{me} Marinette Revenu	Directeur de thèse	Professeur
	Président du jury	

Remerciements

Ce travail de recherche a été réalisé au sein de l'équipe Vision et Analyse d'Images (VAI) du Laboratoire Universitaire des Sciences Appliquées de Cherbourg (LUSAC). Ce travail s'est fait grâce à une collaboration entre le service d'anatomie et cytologie pathologiques de l'hôpital Pasteur de Cherbourg et l'Université de Caen, mais également avec le Groupe de REcherches en Informatique, Image et Instrumentation de Caen (GREYC). Il a été financé par une bourse de la ligue contre le cancer.

Je tiens à remercier Madame Marinette Revenu, Professeur de l'ISMRA, pour avoir dirigé mon doctorat et guidé mes recherches. Je tiens également à remercier Le Docteur Hubert Elie, chef du service d'anatomie et cytologie pathologiques de l'hôpital Pasteur de Cherbourg, pour m'avoir accueilli dans ses locaux et avoir été l'initiateur de mon doctorat.

Je remercie également Monsieur Pierre Bonton, Professeur de l'Université de Clermont-Ferrand, et Monsieur Alain Faure, Professeur de l'Université du Havre, d'avoir accepté d'être rapporteurs de cette thèse. Je leur adresse tous mes remerciements.

J'exprime toute ma gratitude à Hubert Cardot, Maître de conférences de l'Université de Caen, pour m'avoir encadré et avoir su me faire confiance. Je remercie également Monsieur Abderrahim Elmoataz, Maître de conférences de l'Université de Caen, pour son encadrement ainsi que ses conseils avisés de traiteur d'images. Que ces deux personnes, qui se sont investies pour la réussite de ma thèse, voient ici l'expression de toute ma reconnaissance.

Je remercie également les membres du LUSAC et particulièrement Monsieur Gilles Gougeon, Maître de conférences de l'Université de Caen, pour s'être occupé des problèmes matériels. Je remercie les membres du GREYC tant le groupe Image que le groupe Apprentissage. Les personnes qui m'ont soutenues sont nombreuses et je les remercie. Les doctorants et les maîtres de conférences font régner une ambiance chaleureuse. Je remercie également tout le personnel du laboratoire d'anapath de Cherbourg pour m'avoir permis de totalement m'intégrer dans le milieu hospitalier.

Je tiens également à remercier tous mes camarades qu'ils soient de Caen ou d'Avranches, chacun d'entre eux a su favoriser mon ouverture d'esprit pour m'aider à aller de l'avant. Un grand merci à Stéphane Durand, Dominique Fournier, Erwan Millour, Willy Queruel et tous les autres.

Enfin ces remerciements ne seraient pas complets si je ne les associais à ma famille qui m'a soutenu et supporté pendant ces années de thèse.

Je pense également à ma chère et tendre.

SOMMAIRE

Partie A : Segmentation et caractérisation couleur 12

1. LES MODÈLES DE LA VISION HUMAINE ET LES ESPACES COULEUR 17

1.1.	LES MODÈLES DE LA VISION HUMAINE	18
1.1.1.	LE MODÈLE DU TRI-STIMULUS	18
1.1.2.	LE MODÈLE DES COULEURS OPPOSÉES	19
1.1.3.	LE MODÈLE DE FAUGERAS	19
1.1.4.	LE MODÈLE DE PRATT	20
1.2.	LES ESPACES COULEUR	20
1.2.1.	L'ESPACE DE COULEUR RVB	20
1.2.2.	L'ESPACE DE COULEUR XYZ	21
1.2.3.	L'ESPACE DE COULEUR $L^*U^*V^*$	23
1.2.4.	L'ESPACE DE COULEUR $L^*A^*B^*$	23
1.2.5.	L'ESPACE DE COULEUR HSL (HUE, SATURATION, LUMINANCE)	24
1.2.6.	L'ESPACE DE COULEUR DE CHASSERY	25
1.2.7.	L'ESPACE $I_1 I_2 I_3$	26
1.2.8.	AUTRES ESPACES DE COULEUR	26
1.2.9.	LA TRANSFORMATION DE RVB À XYZ	26

2. SEGMENTATION COULEUR 29

2.1.	MÉTHODES DE SEGMENTATION COULEUR	29
2.1.1.	PRINCIPALES MÉTHODES DE SEGMENTATION COULEUR	29
2.1.2.	LA LIGNE DE PARTAGE DES EAUX	33
	2.1.2.1. Définitions de base	34
	2.1.2.2. Définition par immersion	35
	2.1.2.3. La L.P.E contrainte par marqueurs	38
2.1.3.	LA LIGNE DE PARTAGE DES EAUX COULEUR	38
2.2.	NOTRE APPROCHE DE LA SEGMENTATION COULEUR	39
2.2.1.	NOTRE FONCTION D'AGRÉGATION COULEUR	39
2.2.2.	OPTIMISATION DE LA FONCTION D'AGRÉGATION	40
2.2.3.	CHOIX DE L'ESPACE COULEUR	41
	2.2.3.1. Comparaison des distances selon l'espace couleur	42
	2.2.3.2. Un critère de choix de l'espace couleur basé sur le contraste	45
	2.2.3.3. Une représentation optimale de la couleur par apprentissage d'informations a priori	48
	2.2.3.3.1. Fusion de composantes couleur	49
	2.2.3.3.2. Apprentissage par sélection séquentielle de composantes couleur	49
	2.2.3.3.3. Résultats	50
	2.2.3.4. Bilan	51
2.2.4.	GRADIENT COULEUR	52
2.3.	EXPÉRIMENTATIONS DE NOTRE APPROCHE DE SEGMENTATION COULEUR	53
2.3.1.	NOTATIONS	53
2.3.2.	CYTOLOGIE DES SÉREUSES	54
	2.3.2.1. Problématique et description des images	54

2.3.2.2.	Stratégie de segmentation.....	55
2.3.2.3.	Détection des cellules intéressantes.....	56
2.3.2.4.	Elimination des globules rouges.....	57
2.3.2.5.	Extraction des régions cytoplasmiques.....	59
2.3.2.6.	Extraction des régions nucléaires	61
2.3.2.7.	Séparation des noyaux	66
2.3.2.8.	Division des cytoplasmes	67
2.3.3.	NEUDS DANS DES PANNEAUX DE DE BOIS	69
2.3.4.	CONCLUSION SUR LES EXPÉRIMENTATIONS	73

3. CARACTÉRISATION COULEUR **74**

3.1.	INTRODUCTION	74
3.2.	TAILLE	75
3.3.	FORME	76
3.4.	TEXTURE COULEUR	79
3.4.1.	MÉTHODES USUELLES	79
3.4.2.	MESURES STATISTIQUES	80
3.4.3.	LIGNE DE PARTAGE DES EAUX MULTI-ÉCHELLE	80
3.4.4.	TEXTONS	84
3.4.5.	CHAPEAU HAUT DE FORME	88
3.4.6.	L'ESPACE COULEUR EN ANALYSE DE TEXTURE	89

4. CONCLUSION **91**

5. RÉFÉRENCES **92**

Partie B : Classification par réseaux de neurones **98**

1. MONNA : UNE ARCHITECTURE DE RÉSEAUX DE NEURONES **103**

1.1.	INTRODUCTION AUX RÉSEAUX DE NEURONES	103
1.1.1.	LE MODÈLE DE NEURONE FORMEL	103
1.1.2.	LES RÉSEAUX MULTICOUCHES	104
1.1.3.	PROBLÈMES LIÉS AUX RÉSEAUX MULTICOUCHES	107
1.1.3.1.	Structure du réseau	107
1.1.3.2.	Prétraitement des entrées	108
1.2.	PROBLÉMATIQUE	108
1.3.	PRINCIPE DE L'ARCHITECTURE	109
1.4.	CONSTRUCTION D'UNE ARCHITECTURE : MONNA	109
1.5.	CRÉATION DE L'ARCHITECTURE	110
1.6.	APPRENTISSAGE DES RÉSEAUX DE NEURONES	110
1.6.1.	PRÉTRAITEMENT DES DONNÉES	111
1.6.2.	STRUCTURE ADAPTATIVE	111
1.6.3.	DÉTAILS SUR L'APPRENTISSAGE	113
1.7.	HIÉRARCHISATION DES RÉSEAUX DE NEURONES	115
1.7.1.	NOTATIONS	117
1.7.2.	LA SÉLECTION PAR ÉLIMINATION	117

1.8.	EXPÉRIMENTATIONS DE L'ARCHITECTURE MONNA	118
1.8.1.	DESCRIPTION DES BASES DE DONNÉES	118
1.8.1.1.	Base Liver Bupa	118
1.8.1.2.	Base Pima	118
1.8.1.3.	Base Ionosphere.....	118
1.8.1.4.	Base Cancer	118
1.8.1.5.	Base Vehicle	118
1.8.1.6.	Base Pageblocks	119
1.8.1.7.	Base Glass	119
1.8.1.8.	Base Segmentation	119
1.8.1.9.	Base OptDigits.....	119
1.8.1.10.	Base Serous	119
1.8.1.11.	Récapitulatif	120
1.8.2.	COMPARAISON GLOBALE ENTRE UN RMC ET MONNA	120
1.8.3.	LA STRUCTURE ADAPTATIVE	122
1.8.4.	LE NOMBRE DE NEURONES DE LA COUCHE CACHÉE	123
1.8.5.	LA SÉLECTION PAR ÉLIMINATION	125
1.8.6.	COMPARAISON GÉNÉRALE	127
2.	SÉLECTION D'ATTRIBUTS	129
2.1.	LE PROBLÈME DE LA SÉLECTION D'ATTRIBUTS	129
2.2.	DÉFINITION DE LA PERTINENCE	130
2.3.	COMPOSANTES DE BASE DES MÉTHODES DE SÉLECTION DE CARACTÉRISTIQUES	132
2.3.1.	LES ESPACES DE RECHERCHE	135
2.3.2.	LES FONCTIONS D'ÉVALUATION	136
2.4.	APPROCHES FILTER ET WRAPPER	137
2.5.	ALGORITHMES DE RECHERCHE	138
2.5.1.	L'ALGORITHME <i>HILL-CLIMBING</i>	138
2.5.2.	L'ALGORITHME <i>BEST-FIRST SEARCH</i>	139
2.5.3.	L'ALGORITHME <i>SEQUENTIAL FORWARD FLOATING SELECTION (SFFS)</i>	139
2.6.	APPROCHES TESTÉES	140
2.7.	L'APPROCHE FILTER	140
2.7.1.	TRACE DES MATRICES DE COVARIANCE INTER ET INTRA CLASSE	140
2.7.2.	THÉORIE DE L'INFORMATION	141
2.7.3.	MÉTHODE DE YACOUB	142
2.8.	L'APPROCHE WRAPPER	142
2.9.	RÉSULTATS EXPÉRIMENTAUX	142
2.9.1.	COMPARAISON SUR LES TAUX DE RECONNAISSANCE	143
2.9.2.	COMPARAISON DES APPROCHES <i>FILTER</i>	144
2.9.3.	COMPARAISON DES APPROCHES <i>WRAPPER</i> SUR LE NOMBRE D'ATTRIBUTS	147
3.	CONCLUSION	148
4.	RÉFÉRENCES	151

Partie C : Application en microscopie	155
<hr/>	
1. UN SYSTÈME DE TRI CELLULAIRE	158
<hr/>	
1.1. GÉNÉRALITÉS D'ANATOMIE ET DE CYTOLOGIE PATHOLOGIQUES	158
1.1.1. L'ANATOMIE PATHOLOGIQUE	158
1.1.2. HISTORIQUE DE LA CYTOLOGIE	158
1.1.3. STRUCTURE ANATOMIQUE ET CYTOLOGIQUE	159
1.1.4. LES PRÉLÈVEMENTS EN ANATOMIE ET CYTOLOGIE PATHOLOGIQUES	161
1.1.5. LES COLORATIONS	162
1.1.6. TYPES CELLULAIRES EN CYTOLOGIE DES SÉREUSES	162
1.2. LE SCREENING	165
1.2.1. LE SCREENING MANUEL	165
1.2.2. LE SCREENING AUTOMATIQUE	166
1.2.2.1. Historique des systèmes de <i>screening</i>	166
1.2.2.2. Structure d'un système informatique de <i>rescreening</i>	167
1.3. LE SYSTÈME ARCTIC	168
<hr/>	
2. L'ACQUISITION DES IMAGES	170
<hr/>	
2.1. RÉSULTATS EXPÉRIMENTAUX SUR LE SYSTÈME	170
2.1.1. INFLUENCE DU TEMPS DE RÉPONSE CAMÉRA PLUS MICROSCOPE	170
2.1.2. INFLUENCE DE L'ENVIRONNEMENT	172
2.2. DÉTERMINATION DU TYPE DE BRUIT DU SYSTÈME	173
2.2.1. RAPPELS GÉNÉRAUX SUR LE BRUIT	173
2.2.2. BRUIT D'ACQUISITION DE NOTRE SYSTÈME	174
2.3. ELIMINATION DU BRUIT	174
<hr/>	
3. RÉSULTATS EXPÉRIMENTAUX	176
<hr/>	
3.1. EVALUATION DE LA SEGMENTATION	176
3.2. CONSTRUCTION D'UNE BASE DE DONNÉES	178
3.2.1. L'ÉTIQUETAGE PAR UN EXPERT	178
3.2.2. INTERSECTION ENTRE PLUSIEURS EXPERTS	180
3.3. CLASSIFICATION DES OBJETS	181
<hr/>	
4. FONCTIONNALITÉS DU SYSTÈME ARCTIC	183
<hr/>	
5. CONCLUSION	187
<hr/>	
6. RÉFÉRENCES	188
<hr/>	
Conclusion et perspectives	192
<hr/>	

Introduction

Durant ces dernières années, l'informatique s'est affirmée comme une nouvelle technologie majeure. Les progrès des moyens informatiques et des techniques de traitement ont permis d'ouvrir une voie de développement très prometteuse. L'amélioration et la multiplication des moyens de diffusion des données informatiques a pris également une grande part à la multiplication de ces techniques. Les outils de communication utilisent de plus en plus les images, ce qui a entraîné un accroissement des recherches dans les domaines du traitement et de l'analyse d'images. Ces techniques ont été développées avec succès dans des domaines divers et très variés (images industrielles, de robotique, biologiques, aériennes, etc.).

Dans le domaine médical, l'utilisation des images est à présent considérée comme indispensable pour des raisons de facilité de stockage et aussi pour disposer de techniques d'analyse des images automatiques ou semi-automatiques grâce à des systèmes informatiques. Ces systèmes permettent de proposer aux praticiens une nouvelle façon de travailler en obtenant de nouvelles mesures par des techniques quantitatives et objectives. Le challenge est de taille, la réalisation de systèmes informatiques aidant le médecin doit permettre de sécuriser le diagnostic, le rendant ainsi plus fiable et sécurisé. Les conséquences peuvent être nombreuses. L'établissement d'un diagnostic fiable peut permettre par exemple d'améliorer la qualité du dépistage des cancers et donc de favoriser la prévention.

Cette recherche de qualité dans l'établissement d'un diagnostic est très présente en anatomie et cytologie pathologiques. En cytologie, le pathologiste doit établir, à l'issue d'examens minutieux de lames de cellules, un diagnostic qui doit être le plus fiable possible. Le pathologiste reconnaît les types cellulaires présents sur une lame et c'est cette étude de la lame qui détermine son diagnostic. Cet examen se fait au microscope, il est donc visuel, manuel et par conséquent très fastidieux. Afin de faciliter le travail de lecture des lames, des recherches ont permis de développer des systèmes semi-automatiques permettant de sécuriser l'établissement du diagnostic. Le pathologiste est alors assisté dans son travail. De tels systèmes font appel à des techniques de traitement et d'analyse d'images, d'intelligence artificielle et de biologie.

Dans le cadre du pôle *Traitement et Analyse d'Images*, les différents acteurs que sont le *Laboratoire Universitaire des Sciences Appliquées de Cherbourg (LUSAC)*, le *Groupe de REcherches en Informatique, Image et Instrumentation de Caen (GREYC)* ainsi que le service d'anatomie et de cytologie pathologiques de l'hôpital Pasteur de Cherbourg travaillent sur le traitement et l'analyse d'images 2D de cellules. L'élaboration de nos travaux se déroule donc dans un environnement pluridisciplinaire.

Nous cherchons à réaliser un système informatique semi-automatique permettant de réaliser la détection de cellules anormales sur une lame de prélèvement cytologique. L'élaboration d'un tel système doit évidemment se faire en collaboration entre les différents experts. Le système que nous avons développé se nomme *Aide à la Recherche en Cytologie par le Tri Cellulaire (ARCTIC)*. Il est basé sur trois étapes séquentielles :

- ❶ La segmentation : elle consiste à extraire des objets à partir d'images couleur ; ce sont des composantes cellulaires pour notre application.
- ❷ La caractérisation : elle fournit des caractéristiques des objets précédemment extraits. Ces données doivent être suffisamment pertinentes et précises pour décrire les objets.
- ❸ La classification : les objets, i.e. les cellules, sont classés selon un type cellulaire grâce à une méthode de classification.

Le but de notre système est de détecter les cellules anormales ou suspectes qui auraient pu échapper à la vigilance du pathologiste. Nous nous plaçons donc dans le cadre de l'assurance qualité à travers la sécurisation du diagnostic. Notre système traite des images couleur de la cytologie des séreuses sur lesquelles il doit reconnaître les cellules. Ces cellules sont ensuite classées afin de déterminer leur éventuelle malignité. Les cellules suspectes ou anormales sont présentées au pathologiste afin de vérifier son diagnostic, le rendant ainsi plus fiable. Nous avons cherché à réaliser un système le plus complet possible et directement utilisable par des pathologistes. Il est facile d'utilisation et convivial car son développement s'est fait en grande concertation avec ces derniers.

L'objectif de la thèse consiste à modéliser et à développer des outils de traitement et d'analyse d'images pour les étapes de segmentation et de caractérisation, mais également des outils de traitement pour l'étape de classification. Chacune de ces étapes doivent être validées par les experts des différents domaines et ceci par des mesures quantitatives visant à construire un système le plus fiable possible. L'objectif du système est d'obtenir un taux maximum de reconnaissance des cellules anormales. La thèse s'articule autour de trois parties, la première traite de la segmentation et de la caractérisation couleur, la deuxième de la classification par réseaux de neurones et la troisième et dernière partie de l'application proprement dite en microscopie.

Plan de lecture

Dans une première partie, nous exposons les principes liés à la perception de la couleur et en particulier les modèles de perception. Nous détaillons ensuite les différents espaces de couleurs de la littérature en détaillant leurs caractéristiques. Puis, nous nous intéressons au principe de la segmentation d'images couleur. Après un rapide état de l'art, nous focalisons notre étude aux approches provenant de la morphologie mathématique qui nous semblent très appropriées pour la segmentation d'images très complexes telles que les images de cytologie. Nous détaillons notre méthode de segmentation d'images couleur qui est basée sur l'établissement d'une fonction d'agrégation couleur de pixels régissant l'algorithme de la ligne de partage des eaux. Nous abordons ensuite chacun des points importants de la méthode et nous déterminons la façon de les aborder et de les optimiser. Nous nous attardons plus particulièrement sur le choix de l'espace couleur et le choix des marqueurs. Notre propos est illustré à travers deux applications provenant respectivement de cytologie et de l'étude des nœuds présents sur des planches de bois. Pour ces deux applications, nous définissons une stratégie complète de segmentation des objets en utilisant notre méthode de segmentation couleur associée à une stratégie d'optimisation de la fonction d'agrégation. Enfin, nous abordons le problème de la caractérisation d'objets extraits lors de la segmentation. Nous cherchons à déterminer des attributs suffisamment pertinents pour permettre une future bonne discrimination des objets. La caractérisation porte principalement sur l'établissement de critères de texture pour la caractérisation de cellules à partir d'images couleur.

Dans une deuxième partie, nous présentons une architecture de réseaux de neurones utile pour la classification de données présentant un grand nombre de classes. Le principe de cette architecture consiste à simplifier la tâche d'un classifieur en simplifiant le problème à résoudre. Après avoir exposé les limites des réseaux multicouches, nous présentons les fondements de notre architecture. Basé sur un raisonnement de type « diviser pour régner », nous définissons une forêt de réseaux de neurones qui permettent de simplifier le problème de reconnaissance, d'améliorer l'apprentissage, de simplifier la structure des réseaux utilisés, et ainsi d'augmenter le taux de reconnaissance du classifieur. Nous présentons chacun des fondements de notre architecture MONNA (*Multiple Ordinate Neural Network Architecture*) à savoir la création, l'apprentissage et la hiérarchisation. Nous comparons ensuite notre architecture avec un réseau multicouche classique sur différentes bases de données et nous montrons que notre approche permet d'obtenir de meilleurs résultats lorsque le nombre de classes est supérieur à 2. Nous détaillons enfin le problème de la sélection d'attributs pertinents qui permet d'une part de simplifier la structure des réseaux de neurones utilisés mais également d'améliorer le taux de reconnaissance de l'architecture MONNA. Après avoir posé le problème de la sélection d'attributs pertinents, nous décrivons les approches communément utilisées (*filter et wrapper*). Puis nous les comparons en les appliquant à notre architecture MONNA.

Dans une troisième partie, nous présentons le problème du tri cellulaire en anatomie et cytologie pathologiques et plus particulièrement pour la cytologie des séreuses. Après quelques rappels de pathologie, nous présentons le problème à résoudre, à savoir la construction d'un système semi-automatique de *screening*, le *screening* désignant l'acte de repérer des cellules sur une lame de prélèvements cytologiques. Nous détaillons le fonctionnement du système ARCTIC et sa place dans le processus d'évaluation d'une lame.

Puis nous détaillons les conditions expérimentales dans lesquelles nous avons obtenu des images couleur à travers un procédé d'acquisition. Enfin, nous présentons les résultats du système. Celui-ci utilise la segmentation et la caractérisation couleur exposées dans la première partie ainsi que l'architecture MONNA et la sélection d'attributs exposées dans la deuxième partie. Nous présentons pour finir les fonctionnalités du système.

La thèse ayant été réalisée dans un environnement pluridisciplinaire, chacune des trois parties présentées peut s'adresser à des experts provenant de domaines différents. L'expert en vision artificielle sera intéressé par les trois parties, l'expert en traitement d'images le sera par la première, l'expert en classification par la deuxième et le pathologiste par la troisième. Ainsi ces trois parties sont relativement indépendantes bien que certaines utilisent des résultats provenant des parties précédentes. Le lecteur pourra donc intervertir la lecture des parties à sa guise.

Segmentation et caractérisation couleur

« La Couleur est par excellence la partie de l'art qui détient le don magique. Alors que le sujet, la forme, la ligne s'adressent d'abord à la pensée, la couleur n'a aucun sens pour l'intelligence, mais elle a tous les pouvoirs sur la sensibilité, elle remue les sentiments »

Journal d'Eugène Delacroix, Peintre.

1. LES MODÈLES DE LA VISION HUMAINE ET LES ESPACES COULEUR.....	17
1.1. LES MODÈLES DE LA VISION HUMAINE	18
1.1.1. LE MODÈLE DU TRI-STIMULUS	18
1.1.2. LE MODÈLE DES COULEURS OPPOSÉES	19
1.1.3. LE MODÈLE DE FAUGERAS.....	19
1.1.4. LE MODÈLE DE PRATT.....	20
1.2. LES ESPACES COULEUR	20
1.2.1. L'ESPACE DE COULEUR RVB	20
1.2.2. L'ESPACE DE COULEUR XYZ	21
1.2.3. L'ESPACE DE COULEUR $L^*U^*V^*$	23
1.2.4. L'ESPACE DE COULEUR $L^*A^*B^*$	23
1.2.5. L'ESPACE DE COULEUR HSL (HUE, SATURATION, LUMINANCE)	24
1.2.6. L'ESPACE DE COULEUR DE CHASSERY	25
1.2.7. L'ESPACE $I_1 I_2 I_3$	26
1.2.8. AUTRES ESPACES DE COULEUR.....	26
1.2.9. LA TRANSFORMATION DE RVB À XYZ.....	26
2. SEGMENTATION COULEUR.....	29
2.1. MÉTHODES DE SEGMENTATION COULEUR.....	29
2.1.1. PRINCIPALES MÉTHODES DE SEGMENTATION COULEUR.....	29
2.1.2. LA LIGNE DE PARTAGE DES EAUX	33
2.1.2.1. Définitions de base.....	34
2.1.2.2. Définition par immersion	35
2.1.2.3. La L.P.E contrainte par marqueurs	38
2.1.3. LA LIGNE DE PARTAGE DES EAUX COULEUR.....	38
2.2. NOTRE APPROCHE DE LA SEGMENTATION COULEUR.....	39
2.2.1. NOTRE FONCTION D'AGRÉGATION COULEUR.....	39
2.2.2. OPTIMISATION DE LA FONCTION D'AGRÉGATION	40
2.2.3. CHOIX DE L'ESPACE COULEUR	41
2.2.3.1. Comparaison des distances selon l'espace couleur.....	42
2.2.3.2. Un critère de choix de l'espace couleur basé sur le contraste	45
2.2.3.3. Une représentation optimale de la couleur par apprentissage d'informations a priori	49
2.2.3.4. Bilan	51
2.2.4. GRADIENT COULEUR.....	52
2.3. EXPÉRIMENTATIONS DE NOTRE APPROCHE DE SEGMENTATION COULEUR	53
2.3.1. NOTATIONS	53
2.3.2. CYTOLOGIE DES SÉREUSES	54
2.3.2.1. Problématique et description des images	54
2.3.2.2. Stratégie de segmentation	55
2.3.2.3. Détection des cellules intéressantes.....	56
2.3.2.4. Elimination des globules rouges	57
2.3.2.5. Extraction des régions cytoplasmiques.....	59
2.3.2.6. Extraction des régions nucléaires	61
2.3.2.7. Séparation des noyaux	66
2.3.2.8. Division des cytoplasmes.....	67
2.3.3. NŒUDS DANS DES PANNEAUX DE DE BOIS.....	69
2.3.4. CONCLUSION SUR LES EXPÉRIMENTATIONS	73

3. CARACTÉRISATION COULEUR	74
3.1. INTRODUCTION	74
3.2. TAILLE.....	75
3.3. FORME.....	76
3.4. TEXTURE COULEUR.....	79
3.4.1. MÉTHODES USUELLES	79
3.4.2. MESURES STATISTIQUES	80
3.4.3. LIGNE DE PARTAGE DES EAUX MULTI-ÉCHELLE	80
3.4.4. TEXTONS.....	84
3.4.5. CHAPEAU HAUT DE FORME	88
3.4.6. L'ESPACE COULEUR EN ANALYSE DE TEXTURE	89
4. CONCLUSION.....	91
5. RÉFÉRENCES.....	92

Figures

FIGURE 1. LE SPECTRE DES COULEURS VISIBLES [NICOLE 97].....	17
FIGURE 2. LES LARGEURS DE BANDE SPECTRALE DES CÔNES PHOTORÉCEPTEURS [KUNT 91].....	19
FIGURE 3. LE MODÈLE DES COULEURS OPPOSÉES [KUNT 91].....	19
FIGURE 4. L'OBTENTION DES COMPOSANTES DU SYSTÈME DE FAUGERAS [KUNT 91].....	20
FIGURE 5. LA REPRÉSENTATION SPATIALE DU MODÈLE <i>RVB</i> [POYNTON 95].....	21
FIGURE 6. VARIATION DU TRI-STIMULUS <i>RVB</i> EN FONCTION DE LA LONGUEUR D'ONDE [KUNT 91].....	21
FIGURE 7. VARIATION DU TRI-STIMULUS <i>XYZ</i> EN FONCTION DE LA LONGUEUR D'ONDE [KUNT 91].....	21
FIGURE 8. LE DIAGRAMME DE CHROMATICITÉ DE LA CIE. LES VARIABLES DE CHROMINANCE SONT PORTÉES SUR LES AXES [POYNTON 95].....	22
FIGURE 9. L'ESPACE <i>xyY</i> DANS SA REPRÉSENTATION SPATIALE [POYNTON 95].....	23
FIGURE 10. LA REPRÉSENTATION SPATIALE DU MODÈLE $L^*A^*B^*$ [NICOLE 97].....	24
FIGURE 11. LA REPRÉSENTATION SPATIALE DU MODÈLE <i>HSL</i> [NICOLE 97].....	25
FIGURE 12. LE DIAGRAMME CIE AVEC LES SOURCES <i>A,B,C,D,E</i> [TREMEAU 93].....	27
FIGURE 13. MINIMA, BASSINS VERSANTS ET LIGNE DE PARTAGE DES EAUX [VINCENT 90].....	35
FIGURE 14. BARRAGES AUX ENDROITS OÙ L'EAU PROVENANT DE DEUX MINIMA SE MÉLANGERAIENT [VINCENT 90].	36
FIGURE 15. LA DISTANCE GÉODÉSIQUE ENTRE <i>x</i> ET <i>y</i> À L'INTÉRIEUR DE <i>A</i> EST LE PLUS PETIT CHEMIN ENTRE CES DEUX POINTS QUI EST TOTALEMENT INCLUS DANS <i>A</i> [VINCENT 90].....	36
FIGURE 16. LA ZONE D'INFLUENCE GÉODÉSIQUE DE COMPOSANTES CONNECTÉES B_i À L'INTÉRIEUR D'UN ENSEMBLE <i>A</i> [VINCENT 90].....	37
FIGURE 17. ITÉRATION ENTRE X_h ET X_{h+1} [VINCENT 90].....	37
FIGURE 18. L'IMAGE ORIGINALE ET LES IMAGES DE DISTANCE AU BLANC DANS CHAQUE ESPACE COULEUR.....	43
FIGURE 19. HISTOGRAMMES DES FIGURES 18 (B) À (G).....	44
FIGURE 20. LES IMAGES PROVENANT DE CONTEXTES MICROSCOPIQUE, EXTÉRIEUR ET INTÉRIEUR.....	47
FIGURE 21. LES RÉGIONS DES IMAGES DES FIGURE 20(A), (B) (C) ET (D).....	48
FIGURE 22. LES RÉGIONS OBTENUES AVEC L'IMAGE DE COMBINAISON, LA COMPOSANTE BLEUE ET LA COMPOSANTE X_j	51
FIGURE 23. LES DIFFÉRENTES CONFIGURATIONS SPATIALES DES CELLULES.....	55
FIGURE 24. LA STRATÉGIE DE SEGMENTATION D'IMAGES DE CYTOLOGIE DES SÉREUSES.....	56
FIGURE 25. EXEMPLES D'EXTRACTION DES CELLULES INTÉRESSANTES.....	57
FIGURE 26. UNE IMAGE DE CELLULES CONTENANT DES GLOBULES ROUGES ET LE MASQUE CORRESPONDANT.....	58
FIGURE 27. UNE IMAGE DE CELLULES EN AMAS ACCOLÉES AVEC UN DÉBRIS CELLULAIRE ET LE MASQUE CORRESPONDANT.....	58
FIGURE 28. LES MARQUEURS (A), LE GRADIENT (B), LES MARQUEURS ÉTIQUETÉS (C) ET LES RÉGIONS CYTOPLASMIQUES EN JAUNE (D).....	60
FIGURE 29. GRAPHE DE LA SEGMENTATION DES CYTOPLASMES.....	61
FIGURE 30. OBTENTION DE L'IMAGE I_S À PARTIR D'UNE IMAGE COULEUR.....	64
FIGURE 31. LES MARQUEURS (A), LE GRADIENT (B), LES MARQUEURS ÉTIQUETÉS (C) ET LES RÉGIONS NUCLÉAIRES EN JAUNE (D).....	65
FIGURE 32. GRAPHE DE LA SEGMENTATION DES NOYAUX.....	65
FIGURE 33. DES NOYAUX ACCOLÉS (A) ET (B), SONT SÉPARÉS (F) : LA FONCTION DISTANCE (B) VUE 3D DE LA FONCTION DISTANCE (E) LES ÉRODÉS ULTIMES (D).....	67
FIGURE 34. DEUX EXEMPLES DE SÉPARATION DE NOYAUX.....	67
FIGURE 35. IMAGES PRÉSENTANT LA SÉPARATION DES CYTOPLASMES.....	69
FIGURE 36. LES DIFFÉRENTS TYPES DE NŒUDS.....	70
FIGURE 37. LES DIFFÉRENTES IMAGES OBTENUES LORS DE L'EXTRACTION DES MARQUEURS ET DE LA LPE.....	71
FIGURE 38. GRAPHE DE LA SEGMENTATION DES NŒUDS DE BOIS.....	72
FIGURE 39. EXEMPLES DE NŒUDS SEGMENTÉS.....	73
FIGURE 40. DES CELLULES DE LA CYTOLOGIE DES SÉREUSES SUIVANT LEUR CLASSE D'APPARTENANCE.....	75
FIGURE 41 : CARACTÉRISATION DE CELLULES DES SÉREUSES PAR LA SURFACE (A) ET PAR LE PÉRIMÈTRE (B). LES SÉRIES DE NOYAUX CORRESPONDENT AUX SÉRIES VERTICALES DE POINTS.....	76
FIGURE 42. DES OBJETS DE DIFFÉRENTES FORMES (A) ET LEURS ENVELOPPES CONVEXES RESPECTIVES (B).....	77
FIGURE 43. CARACTÉRISATION DES FORMES DE LA FIGURE 42.A.....	78
FIGURE 44. LE POUVOIR DISCRIMINATOIRE D'UN PARAMÈTRE EN FONCTION D'UN AUTRE.....	78
FIGURE 45. MODIFICATION HOMOTOPIQUE DU GRADIENT D'UNE FONCTION D'APRÈS [JACKWAY 95B].....	82
FIGURE 46. LES RÉGIONS OBTENUES PAR LA LIGNE DE PARTAGE DES EAUX MULTI-ÉCHELLE D'UNE CELLULE DES SÉREUSES.....	83

FIGURE 47. LES ÉTAPES DE L'EXTRACTION DE TEXTURE D'UN OBJET DANS L'ESPACE <i>RVB</i>	87
FIGURE 48. LES RÉGIONS OBTENUES PAR LA MÉTHODE DU CHAPEAU HAUT DE FORME DANS L'ESPACE <i>RVB</i>	89
FIGURE 49. LES DIFFÉRENTES MÉTHODES D'EXTRACTION DE TEXTURE DANS PLUSIEURS ESPACES COULEUR AVEC DE GAUCHE À DROITE LES MÉTHODES : TEXTONS, LPE ET CHAPEAU HAUT DE FORME.	90

Tableaux

TABLEAU 1. LES BLANCS DE RÉFÉRENCE DE L'ESPACE <i>XYZ</i> [TREMEAU 93].	27
TABLEAU 2. TABLEAU DES CONTRASTES GLOBAUX ET NOMBRE DE POINTS DE CONTOURS PAR IMAGE.	46
TABLEAU 3. TABLEAU DES CONTRASTES GLOBAUX RAPPORTÉ AU NOMBRE DE POINTS DE CONTOUR PAR IMAGE. ..	47
TABLEAU 4. LES FORMULES DE COMBINAISONS SELON LES IMAGES UTILISÉES DANS LA BASE.	50
TABLEAU 5. LES CONTRASTES SUIVANT LES DIFFÉRENTS FONDS DES IMAGES.	60
TABLEAU 6. CONTRASTES POUR LES DIFFÉRENTES CONFIGURATIONS DE NOYAUX DANS CHAQUE ESPACE COULEUR.63	
TABLEAU 7. LES CONTRASTES POUR PLUSIEURS IMAGES DE NŒUDS.	70
TABLEAU 8. LE NOMBRE DE RÉGIONS SUIVANT LE COEFFICIENT DU LISSAGE SIMULANT L'ASPECT MULTI-ÉCHELLE.83	

L'attrait pour la couleur s'explique en grande partie par le fait que la couleur est l'un des éléments fondamentaux de la perception visuelle. Depuis la préhistoire, la couleur a joué un rôle important dans la communication visuelle. Au fil des siècles les hommes ont commencé à établir les bases d'un art basé sur la couleur. A cette époque cet art était uniquement utilitaire et aucune signification esthétique n'y était attachée. Ce n'est que plus tard que les représentations de formes colorées firent place à l'art véritable et les civilisations des pays du Moyen Orient excellaient dans la pratique de la peinture et de la teinture. Ce large usage de la couleur incita l'homme à rechercher les causes et les raisons de cet aspect des choses si bien que la couleur est restée parmi l'un des phénomènes naturels les plus étudiés des savants, des artistes et des poètes. Dans cette partie, nous allons tout d'abord analyser les modèles de la vision humaine et les espaces couleurs qui existent. Puis, nous étudierons les principes de la segmentation couleur et plus particulièrement ceux provenant de la morphologie mathématique. Nous exposerons une méthode de segmentation couleur basée sur la ligne de partage des eaux couleur. Quelques applications seront présentées afin de valider la méthode. Enfin, nous verrons des méthodes de caractérisation couleur des objets extraits lors de la segmentation.

1. Les modèles de la vision humaine et les espaces couleur

La première tentative de compréhension du phénomène de la couleur fut entreprise par le philosophe Aristote qui étudia, au quatrième siècle avant notre ère, le mélange des couleurs en observant les effets de différents verres colorés sur la lumière. Au dix-septième siècle le physicien britannique Isaac Newton développa la théorie spectrale de la lumière en dirigeant un faisceau de lumière solaire sur la face d'un prisme. En 1801, le médecin anglais Thomas Young établit la théorie trichromatique de la lumière qui fût améliorée et complétée par Hermann Von Helmholtz, cette théorie porte d'ailleurs le nom de loi Young-Helmholtz. Cette loi stipule que la rétine doit posséder des éléments de trois espèces sensibles respectivement au bleu-indigo, au vert-jaune et au rouge afin de posséder la vision totale des couleurs.

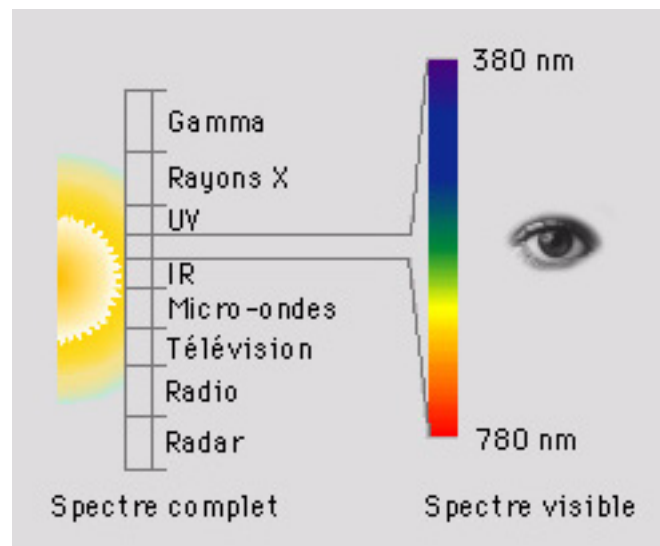


Figure 1. Le spectre des couleurs visibles [Nicole 97].

Tous ces travaux ont permis de définir plus précisément la couleur. Ce que nous appelons communément la couleur est le résultat de la perception d'ondes de lumière sur la rétine dans une région visible du spectre, à savoir les longueurs d'onde comprises entre $380nm$

($1 \text{ nm} = 10^{-9} \text{ m}$) pour le violet et 780 nm pour le rouge (Figure 1). Le spectre des couleurs visibles est continu, il n'existe donc pas de couleur pouvant être véritablement appelée rouge, verte ou bleue car les transitions dans le spectre sont douces et produisent un effet de fondu (visible dans un arc-en-ciel) [Wiszecki 82].

Une question se pose alors : Comment représenter les différentes teintes du spectre lumineux ? Il existe principalement deux réponses à cette question qui sont la représentation fréquentielle et l'espace des couleurs [Kunt 91]. La première possibilité consiste à décrire un stimulus lumineux S par une estimation de sa composition spectrale. Pour cela on filtre le stimulus à l'aide d'un ensemble de N filtres, le problème posé par cette représentation est évidemment lié au très grand nombre de capteurs nécessaires pour avoir une bonne estimation de la composition spectrale. Une autre possibilité consiste à utiliser un espace couleur à plusieurs dimensions. Ceci porte un nom : c'est la colorimétrie et cela consiste à donner suffisamment de composantes numériques pour décrire une couleur. Ce modèle a été étudié et modélisé par la Commission Internationale de l'Eclairage (CIE). C'est ce type de modèle qui est généralement utilisé comme instrument de représentation de la couleur étant donné sa simplicité.

En 1931 la CIE a défini et adopté des illuminants et observateurs standard afin de représenter l'ensemble des couleurs [CIE 31]. La CIE a ensuite montré par l'expérience psychovisuelle d'égalisation des couleurs qu'en combinant trois stimuli de longueurs d'ondes particulières on pouvait synthétiser presque toutes les couleurs existantes. On comprend mieux alors pourquoi ce type de représentation a été retenu au détriment de la représentation fréquentielle. Comme la CIE a prouvé qu'il suffisait de trois variables indépendantes pour décrire presque toutes les couleurs existantes, les espaces couleurs sont donc des systèmes de coordonnées en trois dimensions où la couleur à analyser a été transformée en un tri-stimulus particulier.

Nous allons étudier tout d'abord les modèles de la vision humaine puis nous verrons les espaces couleur que l'on retrouve principalement dans la littérature.

1.1. Les modèles de la vision humaine

1.1.1. Le modèle du tri-stimulus

La rétine humaine possède trois types de cônes photorécepteurs couleur qui répondent à la lumière incidente avec trois courbes de réponse. Ces réponses sont réparties suivant trois canaux qui sont le rouge, le jaune-vert et le bleu plutôt différenciés par leur largeur de bande spectrale (L : large, M : moyenne, S : petite) : Figure 2.

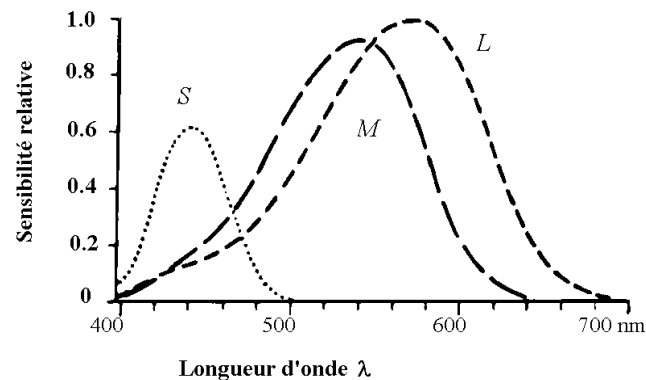


Figure 2. Les largeurs de bande spectrale des cônes photorécepteurs [Kunt 91].

Avec l'expérience de l'égalisation des couleurs par la synthèse additive réalisée par la CIE, on sait qu'en utilisant les trois couleurs rouge, verte et bleue on peut synthétiser presque toutes les couleurs visibles. Les codages des couleurs dans les différents espaces sont principalement basés sur ce modèle qui est le plus répandu.

1.1.2. Le modèle des couleurs opposées

Ce modèle utilise :

- Une composante chromatique qui décrit la réponse de cellules spectralement opposées (jaune-bleu),
- Une autre composante chromatique qui décrit la réponse de cellules spectralement opposées (rouge-vert),
- Une composante achromatique qui décrit la réponse de cellules qui ne sont sensibles qu'aux variations d'intensité du stimulus blanc-noir.

Ces trois composantes peuvent être formées à partir des composantes du tri-stimulus (Figure 3).

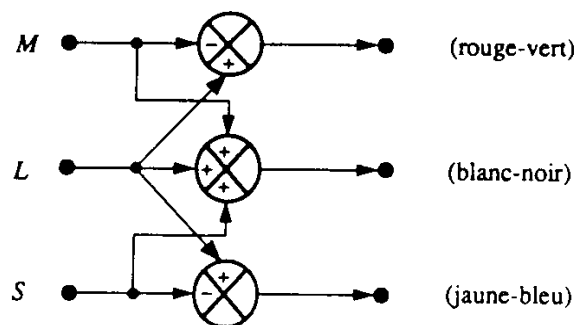


Figure 3. Le modèle des couleurs opposées [Kunt 91].

1.1.3. Le modèle de Faugeras

Ce modèle combine le modèle du tri-stimulus et celui des couleurs opposées en s'appuyant sur le fait que chacun représente un niveau d'abstraction différent [Faugeras 79].

Il est composé de trois parties, chacune de ces parties correspond à un traitement du stimulus lumineux à un niveau donné du système visuel humain. Dans un premier temps la lumière incidente est transformée en trois signaux par les trois types de cônes L , M et S . Ensuite intervient un opérateur logarithmique car des études physiologiques ont montré que la

réponse spectrale des trois types de cônes était proportionnelle au logarithme du stimulus. La troisième étape consiste en la transformation du tri-stimulus en trois composantes : une achromatique A et deux chromatiques C_1 et C_2 (Figure 4).

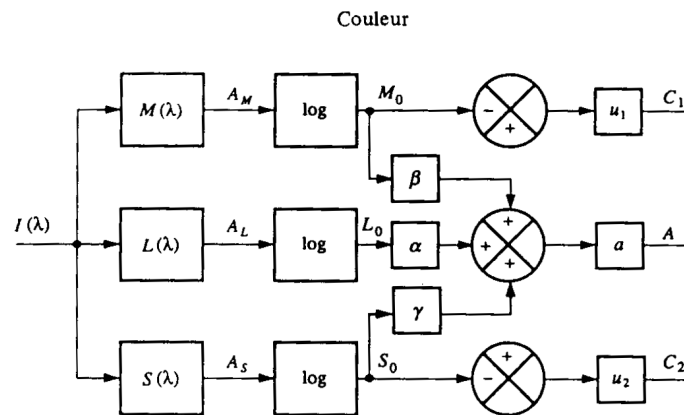


Figure 4. L'obtention des composantes du système de Faugeras [Kunt 91].

1.1.4. Le modèle de Pratt

Bien que des modèles tels que le tri-stimulus soient communément utilisés, il est difficile pour une personne de faire le lien entre la perception de l'apparence visuelle d'une couleur et les composantes dans un espace de couleur [Pratt 78]. Il est plus intuitif d'utiliser les termes de teinte T , de saturation S et de luminance L pour décrire une couleur. Selon la CIE la teinte est la sensation selon laquelle une zone apparaît similaire à une des couleurs perçues : rouge, jaune-vert, bleu ou une combinaison de deux d'entre elles. La saturation est donnée par un adjectif comme vif, terne ou éclatant et est la proportion de couleur par rapport à la luminance. Enfin la luminance est la sensation selon laquelle une zone émet plus ou moins de lumière et est donnée par un adjectif comme clair ou foncé.

L'intérêt d'un espace couleur basé sur ces composantes réside dans leur interprétation facile et intuitive [Levkowitz 93].

1.2. Les espaces couleur

1.2.1. L'espace de couleur RVB

Cet espace est basé sur le modèle du tri-stimulus. C'est un système additif qui décompose les couleurs en trois quantités des trois couleurs primaires : le rouge, le vert et le bleu. C'est le système le plus communément utilisé dans les images couleur et les moniteurs.

Le modèle *RVB* utilise le système de coordonnées cartésien. La diagonale du blanc $(1,1,1)$ au noir $(0,0,0)$ représente les niveaux de gris (Figure 5).

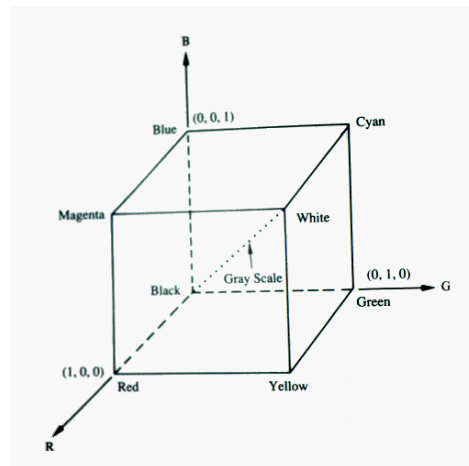


Figure 5. La représentation spatiale du modèle *RVB* [Poynton 95].

1.2.2. L'espace de couleur XYZ

L'espace *XYZ* dépend linéairement du système *RVB* et a été introduit par la CIE pour remédier aux inconvénients du système *RVB* pour lequel certaines couleurs sont représentées par une valeur du stimulus négative (Figure 6).

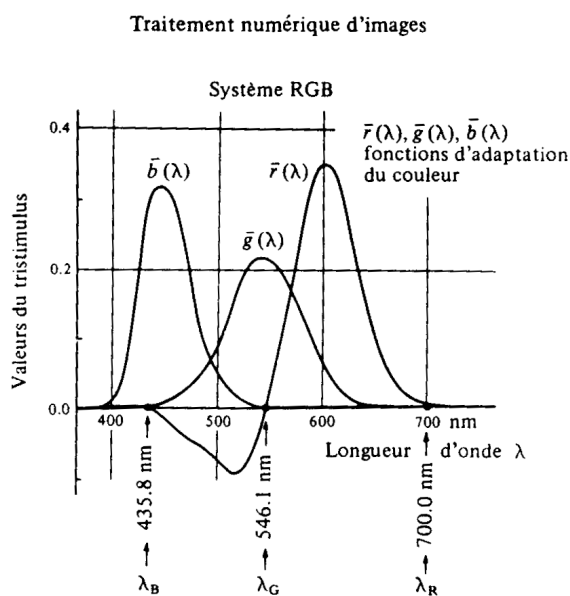


Figure 6. Variation du tri-stimulus *RVB* en fonction de la longueur d'onde [Kunt 91].

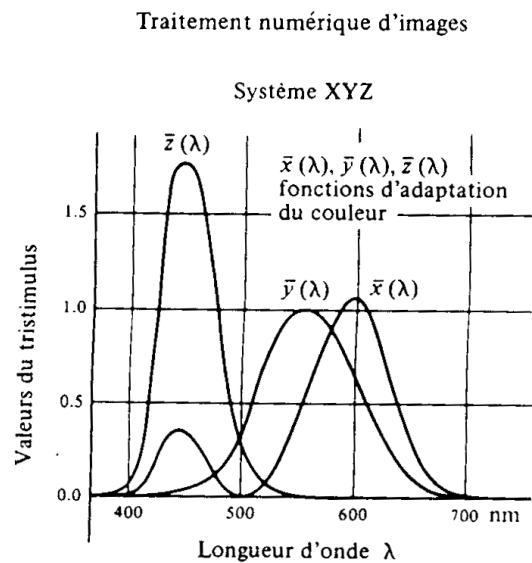


Figure 7. Variation du tri-stimulus *XYZ* en fonction de la longueur d'onde [Kunt 91].

Le système *XYZ* pallie cet inconvénient et une couleur est représentée par une combinaison positive de *X*, *Y* et *Z* (Figure 7).

On a alors

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} x_r & x_v & x_b \\ y_r & y_v & y_b \\ z_r & z_v & z_b \end{pmatrix} * \begin{pmatrix} R \\ V \\ B \end{pmatrix}$$

La matrice de changement de repère est déterminée en mesurant les coordonnées XYZ correspondant aux trois couleurs primaires RVB suivantes : $(R_{max}, 0, 0)$, $(0, V_{max}, 0)$, $(0, 0, B_{max})$ obtenues en appliquant le maximum des valeurs individuellement à chaque stimulus par rapport au blanc de référence utilisé (cf. § 1.2.9).

Y est appelée la luminance (le degré de noir d'une couleur). En normalisant X , Y et Z on obtient le système xyz défini par les relations suivantes :

$$x = \left(\frac{X}{X + Y + Z} \right)$$

$$y = \left(\frac{Y}{X + Y + Z} \right)$$

$$z = \left(\frac{Z}{X + Y + Z} \right)$$

$$x + y + z = 1$$

Comme $x + y + z = 1$, seules les variables x et y sont indépendantes, ces deux variables sont les composantes chromatiques de la couleur. On peut disposer alors du diagramme de chromaticité de la CIE qui décrit le plan défini par les variables de chrominance x et y (Figure 8). Le blanc est représenté par le point définissant un mélange en quantité égale des trois primaires XYZ [Luong 90].

En combinant ce diagramme avec la luminance on obtient l'espace de couleur xyY (Figure 9).

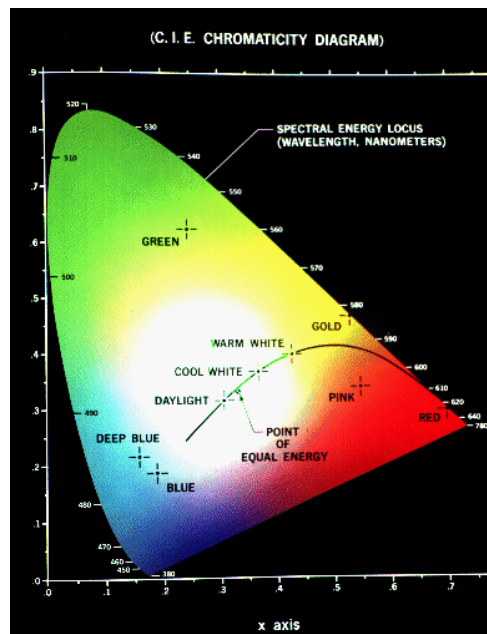


Figure 8. Le diagramme de chromaticité de la CIE. Les variables de chrominance sont portées sur les axes [Poynton 95].

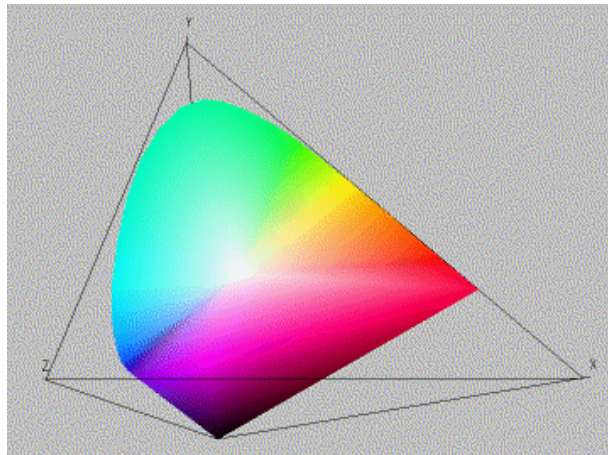


Figure 9. L'espace xyY dans sa représentation spatiale [Poynton 95].

1.2.3. L'espace de couleur $L^*u^*v^*$

Le problème avec l'espace XYZ est son manque d'uniformité dans la perception à savoir que deux couleurs à la même distance d'une autre ne sont pas nécessairement perceptuellement équidistantes [Healey 96]. Pour corriger cet inconvénient la CIE a proposé en 1976 le modèle $L^*u^*v^*$ qui est perceptuellement uniforme, c'est-à-dire que si deux couleurs A et B sont aussi proches entre elles que deux autres couleurs C et D , alors la différence perçue entre A et B sera équivalente à celle perçue entre C et D . Les équations permettant de passer du système XYZ au $L^*u^*v^*$ sont [Tremeau 93] :

$$L^* = 116 \left(\frac{Y}{Y_0} \right)^{\frac{1}{3}} - 16 \quad \text{si} \left(\frac{Y}{Y_0} \right) > 0.008856$$

$$L^* = 903.3 \left(\frac{Y}{Y_0} \right) \quad \text{ailleurs}$$

$$u^* = 13L^* (u' - u'_0)$$

$$v^* = 13L^* (v' - v'_0)$$

$$u' = \frac{4X}{X + 15Y + 3Z} \quad v' = \frac{9X}{X + 15Y + 3Z}$$

$$u'_0 = \frac{4X_0}{X_0 + 15Y_0 + 3Z_0} \quad v'_0 = \frac{9X_0}{X_0 + 15Y_0 + 3Z_0}$$

où X_0 , Y_0 , Z_0 sont les composantes d'un blanc de référence dans l'espace XYZ . L^* est la luminance et u^* et v^* donnent la chrominance de la couleur.

1.2.4. L'espace de couleur $L^*a^*b^*$

L'espace $L^*a^*b^*$ est basé sur le modèle des couleurs opposées.

La variable L^* mesure la luminance (l'opposition noir-blanc) par une valeur entre 0 (noir) et 100 (blanc),

La variable a^* mesure l'opposition rouge-vert par une valeur entre -100 et +100, a^* est positif si la couleur contient du rouge, négatif si la couleur contient du vert et nulle si aucun des deux,

La variable b^* mesure l'opposition jaune-bleu par une valeur entre -100 et +100, b^* est positif si la couleur contient du jaune, négatif si la couleur contient du bleu et nulle si aucun des deux,

Comme pour le système $L^*u^*v^*$, on obtient les valeurs des composantes à partir des valeurs XYZ de la couleur et d'un blanc de référence. Les équations de changement d'espace de XYZ vers $L^*a^*b^*$ sont [Tremeau 93, Pauli 76] :

$$L^* = 116 \left(\frac{Y}{Y_0} \right)^{\frac{1}{3}} - 16 \text{ si } \left(\frac{Y}{Y_0} \right) > 0.008856 \text{ et } L^* = 903.3 \left(\frac{Y}{Y_0} \right) \text{ sinon}$$

$$a^* = 500 \left[f \left(\frac{X}{X_0} \right) - f \left(\frac{Y}{Y_0} \right) \right] \text{ et } b^* = 200 \left[f \left(\frac{Y}{Y_0} \right) - f \left(\frac{Z}{Z_0} \right) \right]$$

avec $f(x) = (x)^{\frac{1}{3}}$ si $x > 0.008856$ et $f(x) = 7.787(x) + \frac{16}{116}$ sinon

Comme le système $L^*u^*v^*$, ce système est perceptuellement uniforme. Sa représentation spatiale est donnée par une sphère (Figure 10).

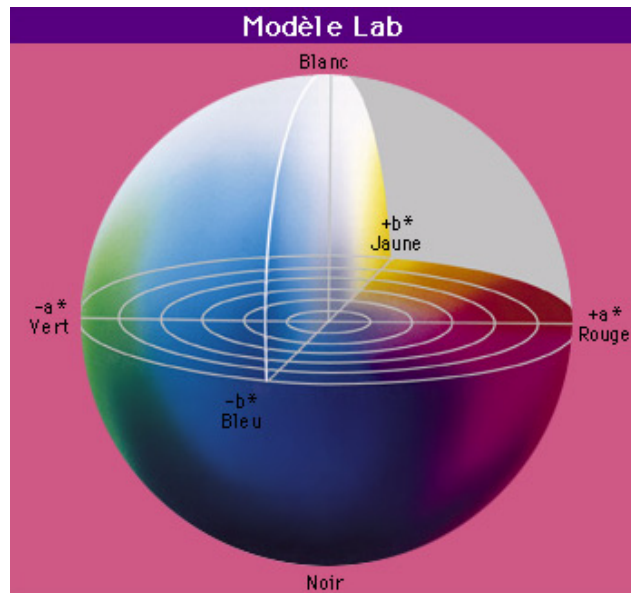


Figure 10. La représentation spatiale du modèle $L^*a^*b^*$ [Nicole 97].

1.2.5. L'espace de couleur HSL (Hue, Saturation, Luminance)

Cet espace est basé sur le modèle de Pratt. Comme son nom l'indique il est basé sur les trois grandeurs de Teinte, Saturation et Luminance. La Figure 11 donne une représentation de cet espace (dans l'espace euclidien RVB). Cet espace est obtenu à partir de l'espace RVB et est décrit par les équations suivantes [Kunt 91, Poynton 95] :

$$H = a \cos \left[\frac{((R-V) + (R-B))}{2\sqrt{(R-V)^2 + (R-B)(V-B)}} \right] \text{ et si } V < B : H = 2\pi - H$$

$$S = 1 - \frac{3 \min(R, V, B)}{L}$$

$$L = R + V + B$$

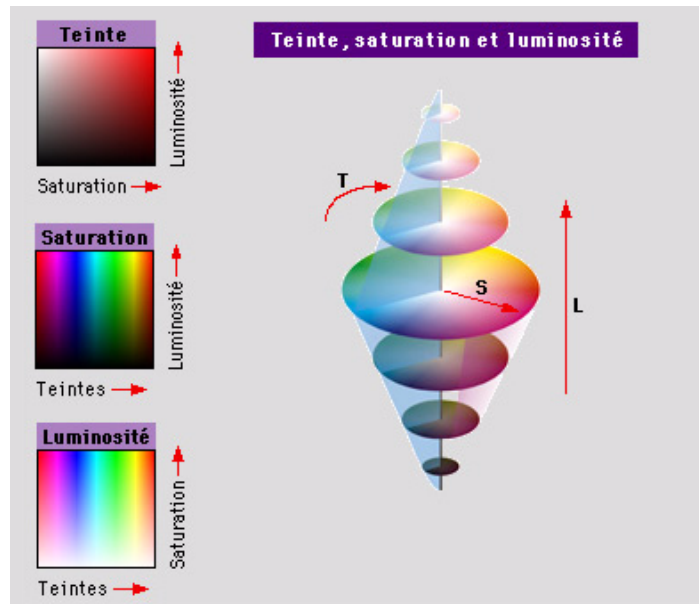


Figure 11. La représentation spatiale du modèle *HSL* [Nicole 97].

La non linéarité des deux équations permettant de calculer H et S entraîne des discontinuités dans leurs valeurs et deux couleurs perceptuellement proches peuvent se transformer en des composantes de teinte et de saturation très différentes. Ceci est dû aux singularités de la teinte pour les points achromatiques ($R=V=B$) et de la saturation pour le noir ($R=V=B=0$).

1.2.6. L'espace de couleur de Chassery

L'espace de Chassery est basé sur le modèle de Faugeras [Faugeras 79] et J.M. Chassery propose un modèle simplifié [Chassery 84] basé sur les composantes :

$$A = \frac{\log(R) + \log(V) + \log(B)}{3}$$

$$C_1 = \sqrt{\frac{3}{2}}(\log(R) - \log(V))$$

$$C_2 = \log(B) - \frac{1}{2}(\log(R) + \log(V))$$

A , C_1 , C_2 correspondent respectivement aux trois composantes du modèle de Faugeras (une achromatique A et deux chromatiques C_1 et C_2). On peut ensuite calculer la saturation et la teinte données par :

$$s = \sqrt{C_1^2 + C_2^2}$$

$$t = a \cos\left(\frac{C_1}{s}\right)$$

1.2.7. L'espace $I_1 I_2 I_3$

Il est basé sur le modèle du tri-stimulus. Il a été obtenu par Ohta [Ohta 80] au moyen d'une transformation de Karhunen-Loève. Celle-ci est un outil statistique qui permet de trouver pour chaque image le système de coordonnées le plus discriminant pour celle-ci. Les composantes X_1, X_2, X_3 calculés ainsi sont :

$$X_i = w_{R_i} R + w_{V_i} V + w_{B_i} B$$

où $W_i = (w_{R_i}, w_{V_i}, w_{B_i})^T$ est le vecteur propre normalisé de la matrice de covariance des distributions de R, V, B correspondant à la valeur propre λ_i avec $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Les composantes obtenues sont non corrélées, X_1 est le vecteur de base optimal (ayant la plus grande variance). Les composantes $I_1 I_2 I_3$ sont donc importantes pour représenter l'information couleur. L'espace $I_1 I_2 I_3$ s'obtient à partir de l'espace RVB par :

$$I_1 = \frac{(R + V + B)}{3} \quad I_2 = \frac{(R - B)}{2} \quad I_3 = \frac{(2V - R - B)}{4}$$

1.2.8. Autres espaces de couleur

Il existe d'autres systèmes d'espaces couleur ($YIQ, CMYK$, etc.). Nous n'avons cité que les espaces principaux utilisés en traitement et analyse d'images. Chacun présente des avantages liés principalement à leur représentation. Il est difficile de prédire a priori comment une méthode va se comporter dans un espace couleur, si bien que plusieurs espaces couleur ont été conçus afin d'aborder de façon différente les problèmes rencontrés dans les applications [CIE 86].

1.2.9. La transformation de RVB à XYZ

Nous avons vu au §1.2.2 que pour passer du système RVB au système XYZ , il faut utiliser une matrice de changement de repère. Dans plusieurs publications et plus particulièrement en analyse d'images couleur, les données relatives aux transformations de RVB à XYZ sont soit erronées, soit incomplètes. Beaucoup de publications omettent de préciser dans quel contexte se place leur transformation (couleurs primaires considérées et blanc de référence retenu). Pour pallier ces inexactitudes, précisons quelques notions indispensables.

Si l'on analyse le domaine des couleurs reproductibles par synthèse additive, nous pouvons constater qu'en fonction des couleurs primaires considérées, ce domaine ne couvre pas la même gamme de couleurs. Il convient donc de préciser quelles sont les couleurs primaires considérées afin d'établir la transformation adéquate.

Cette transformation doit également tenir compte du blanc de référence utilisé pour calibrer les trois couleurs RVB . En fonction de la nature de ce blanc de référence la matrice de transformation est modifiée, dès lors il convient de spécifier quel est le blanc de référence utilisé (Figure 12 et Tableau 1).

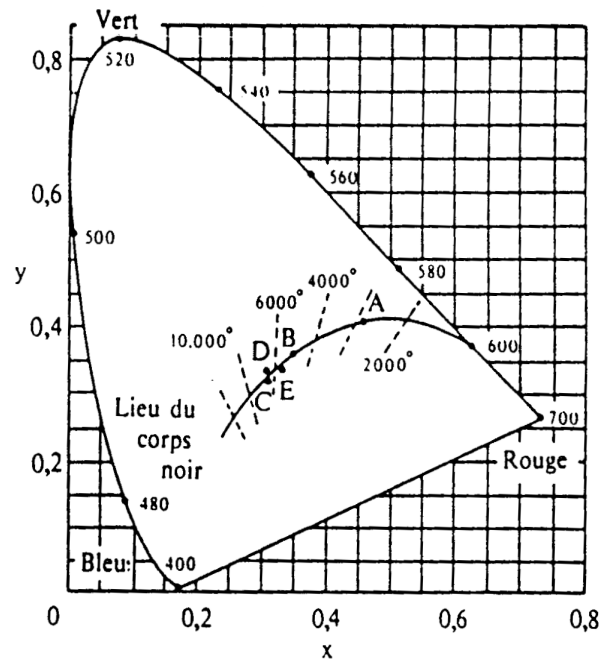


Figure 12. Le diagramme CIE avec les sources A,B,C,D,E [Tremeau 93].

Blanc de référence	X_0	Y_0	Z_0
CIE ILL C	0,9804	1,0	1,1810
CIE ILL D_{65}	0,9502	1,0	1,0881
CIE ILL A	1,0983	1,0	0,3555
CIE ILL E	1,00	1,0	1,0

Tableau 1. Les blancs de référence de l'espace XYZ [Tremeau 93].

La transformation matricielle permettant de passer de *RVB* à *XYZ* doit donc être exprimée par rapport à un blanc de référence et par rapport aux trois stimuli de couleur *RVB* de référence appelés communément primaires. Nous donnons ci-dessous une liste des transformations de base avec leurs éléments de référence associés [Tremeau 93, Wyszecki 82].

Illuminant blanc : **A**

Couleurs de référence : **Macbeth color chart**

$$\begin{array}{lll}
 X_r = 1,083 * 0,996 & X_v = -0,909 * 0,032 & X_b = 0,361 * 0,135 \\
 Y_r = 1,083 * 0,353 & Y_v = 0,909 * 0,701 & Y_b = -0,361 * 0,054 \\
 Z_r = 1,083 * 0,003 & Z_v = -0,909 * 0,097 & Z_b = 0,361 * 1,130
 \end{array}$$

Illuminant blanc : **A**

Couleurs de référence : **Primaires XYZ CIE 1931**

$$\begin{array}{lll}
 X_r = 0,892 & X_v = 0,330 & X_b = 0,083 \\
 Y_r = 0,322 & Y_v = 0,863 & Y_b = 0,004 \\
 Z_r = 0,000 & Z_v = 0,011 & Z_b = 0,409
 \end{array}$$

Illuminant blanc : **C**

Couleurs de référence : **Standard NTSC**

$X_r=0,607$	$X_v=0,174$	$X_b=0,200$
$Y_r=0,299$	$Y_v=0,587$	$Y_b=0,114$
$Z_r=0,000$	$Y_v=0,066$	$Y_b=1,116$

Illuminant blanc : **C**

Couleurs de référence : **Primaires XYZ CIE 1931**

$X_r=0,166$	$X_v=0,125$	$X_b=0,093$
$Y_r=0,060$	$Y_v=0,327$	$Y_b=0,005$
$Z_r=0,000$	$Y_v=0,004$	$Y_b=0,460$

Illuminant blanc : **D₆₅₀₀**

Couleurs de référence : **Visu couleur**

$X_r=0,478$	$X_v=0,299$	$X_b=0,175$
$Y_r=0,263$	$Y_v=0,655$	$Y_b=0,081$
$Z_r=0,020$	$Y_v=0,160$	$Y_b=0,908$

Illuminant blanc : **E**

Couleurs de référence : **Primaires XYZ CIE 1931**

$X_r=2,769$	$X_v=1,7518$	$X_b=1,1300$
$Y_r=1,000$	$Y_v=4,5907$	$Y_b=0,0601$
$Z_r=0,000$	$Y_v=0,0565$	$Y_b=5,5943$

2. Segmentation couleur

Dans cette section nous nous intéressons au problème de la segmentation d'images couleur. Tout d'abord nous présentons les principales méthodes que l'on peut trouver dans la littérature et nous détaillons plus précisément la segmentation par la ligne de partage des eaux couleur. Nous proposons une approche de la segmentation d'images couleur par la ligne de partage des eaux en utilisant une nouvelle fonction d'agrégation couleur. Enfin nous verrons que notre méthode de segmentation peut s'appliquer à différentes classes d'images.

2.1. Méthodes de segmentation couleur

L'emploi de la couleur en segmentation d'images est un sujet de recherche relativement récent. L'évolution de la puissance et de la capacité de stockage des ordinateurs a fourni un regain d'intérêt à l'utilisation d'images couleur depuis quelques années. Bien que l'on trouve plusieurs sortes d'algorithmes de segmentation couleur, la littérature n'est pas aussi riche que pour les images en niveaux de gris. Pourtant les chercheurs ont pensé assez tôt à utiliser la couleur pour segmenter des images. En effet, la couleur étant une information que l'on attache directement et naturellement aux régions, une segmentation couleur devrait permettre d'obtenir des régions qui soient davantage significatives. La couleur a bien sûr été aussi employée pour enrichir des descriptions sémantiques et son intérêt à ce niveau est très élevé mais pour ce qui est de la segmentation couleur les questions que nous nous posons portent uniquement sur son rôle au niveau de l'image. Ce rôle n'est pas aussi évident qu'il pourrait le sembler à la vue d'une image et la segmentation couleur pose des problèmes relatifs au traitement de l'information couleur contenue dans les images. Dans cette section nous ne traiterons que des méthodes de segmentation dites de vision universelle (de bas niveau) et pour lesquelles l'apport de la couleur consiste en un apport quantitatif (par un vecteur à trois dimensions).

2.1.1. Principales méthodes de segmentation couleur

Avant de rappeler les différentes méthodes de segmentation couleur que l'on peut trouver dans la littérature, il nous semble important de préciser en quoi celles-ci peuvent varier. Les méthodes de segmentation couleur présentent plusieurs particularités et que l'on retrouve systématiquement :

1. La méthode utilisée : elle caractérise la manière d'extraire les régions d'une image grâce à son information couleur. Ces méthodes sont variées et nous les verrons par la suite.
2. L'espace couleur utilisé : c'est le choix de cet espace couleur qui pose la principale difficulté pour le traitement des images couleur. Chaque espace possède ses propres caractéristiques et peut présenter un intérêt particulier pour un type d'application. Généralement les méthodes proposées dans la littérature se veulent fonctionner quelque soit l'espace considéré, mais aucun test approprié n'est généralement donné pour choisir l'espace couleur adéquate en fonction de la méthode de segmentation. En ce sens que ce choix ne semble pouvoir être fait qu'une fois la méthode de segmentation mise en place. Plusieurs travaux se sont attachés à comparer une méthode dans différents espaces couleur [Ohta 80, Liu 94, Umbaugh 93], mais le choix de l'espace couleur n'en est pas facilité pour autant. D'autres travaux contournent la difficulté en choisissant d'utiliser plusieurs

espaces couleur en même temps. Le principal problème réside donc dans le choix de l'espace couleur pour la résolution d'une application donnée et en fonction de la méthode de segmentation employée.

3. La métrique utilisée : La distance euclidienne et la distance de Mahalanobis sont les plus communément utilisées. Le problème de la mesure des différences colorimétriques est un des plus importants problèmes qui se pose pour l'analyse de la distribution des couleurs. Evidemment une métrique particulière peut-être utilisée, ceci dépend de l'espace couleur utilisé et de la méthode de segmentation. Par exemple Schettini [Schettini 93] utilise la distance de Fisher pour définir une mesure de similarité entre des régions couleur adjacentes, Treméau [Treméau 97] utilise la distance de Mahalanobis afin de prendre en compte la dispersion des couleurs autour de la couleur moyenne. On notera aussi la distance de Stiles [Lasserre 97, Wyszeccki 82] qui a l'avantage d'être fidèle à la perception des couleurs. On retrouve de façon similaire ce problème de choix de la métrique en niveaux de gris, mais en couleur ce problème est directement lié à la mesure de la perception des couleurs. Pour résumer, le choix de la métrique se fait d'une part, suivant la précision que l'on souhaite atteindre dans la mesure des différences colorimétriques et d'autre part suivant le respect (en terme de distance) de la perception des couleurs.

Nous voyons donc qu'outre la méthode utilisée pour segmenter une image couleur, la principale difficulté (ou plutôt le principal attrait) réside dans le choix de l'espace couleur à utiliser. Voyons les approches les plus communément utilisées pour segmenter une image couleur. Elles se répartissent en cinq groupes :

- ✓ Le clustering,
 - ✓ La croissance de régions,
 - ✓ La morphologie mathématique,
 - ✓ La classification de pixels,
 - ✓ Les autres méthodes.
- ✓ Le clustering

Le clustering consiste à partitionner l'image à partir de projections de l'histogramme de l'image couleur sur des plans ou des droites. Celenk [Celenk 90] utilise des histogrammes $1D$ de composantes couleur calculées dans l'espace $L^*a^*b^*$ afin de détecter par des seuillages les regroupements de pixels de couleurs équivalentes (color clusters). Il utilise ensuite la méthode discriminante linéaire de Fisher pour projeter simultanément les regroupements de couleur sur une ligne, ceci permet de seuiller un histogramme $1D$ et de classer les regroupements de pixels. Robinson [Robinson 98] utilise un histogramme $3D$ de l'image couleur dans l'espace RVB . Cet histogramme est calculé pour un nombre fixe et restreint de couleurs. A partir de différentes couleurs de référence définies pour chacune des régions à isoler, un partitionnement de l'image est calculé par mesure de similarité. On consultera également [Altunbasak 98, Pappas 92]. Les méthodes de seuillage font également partie des méthodes de clustering. On trouve différentes façons d'utiliser l'information couleur pour seuiller une image. Des méthodes itératives qui utilisent plusieurs composantes d'espaces couleur différents [Ohta 80], l'image est progressivement segmentée en régions. L'information couleur est utilisée uniquement sur une composante, qu'elle soit chromatique ou achromatique.

D'autres travaux utilisent simplement une seule composante couleur pour segmenter l'image. Dans ce cas le choix de la composante couleur à utiliser est très déterminant et demande d'étudier au préalable la classe des images couleur à segmenter. Le seuillage peut également s'effectuer sur les trois composantes d'un espace couleur donné, les seuils sont déterminés par une analyse multivariable de la distribution des histogrammes [Garbay 86] ou bien directement par une analyse de l'histogramme 3D [Soille 96].

✓ La croissance de régions

Les méthodes utilisant la croissance de régions sont assez peu nombreuses en segmentation d'images couleur. Nous allons voir quelques unes des ces méthodes et leurs particularités.

- Chassery [Chassery 84] utilise des critères de forme et de couleur pour extraire les régions d'une image couleur. Les régions à isoler sont des cellules qui sont caractérisées par leur couleur et leur forme convexe. Le critère couleur utilisé est une mesure de similarité couleur entre le pixel à ajouter à une région et la région candidate à la croissance. La mesure de différence colorimétrique est calculée dans l'espace de Chassery. Un pixel est ensuite agrégé à une région si la différence de couleur avec la région candidate est inférieure à un certain seuil (fixé ad hoc). Le critère de convexité permet de n'agréger à une région que les points se situant sur l'enveloppe convexe. La croissance de région est initialisée par un seuillage de la composante Verte de l'image couleur *RVB* originale.
- Zugaj [Zugaj 98] utilise une approche coopérative régions-frontières afin d'extraire les informations pertinentes contenues dans une image couleur. Dans un premier temps, un opérateur de gradient couleur particulier (opérant sur des fenêtres de taille fixe) est calculé sur l'image couleur. L'image gradient est ensuite seuillée, les pixels de l'image résultante sont fusionnés entre eux dans un voisinage de huit pixels. Le critère de fusion est une mesure de similarité couleur entre les pixels (fixée ad hoc par un seuil). Une fermeture de contour est calculée afin de fermer les segments disloqués par une recherche du chemin optimal se trouvant à la frontière d'une région couleur de l'image. Les régions finales obtenues sont les régions de couleur homogènes. Tous les calculs sont réalisés dans l'espace *RVB*.
- Tremeau [Tremeau 97] utilise une approche d'extraction et de fusion des régions suivant leurs caractéristiques couleur. Il utilise trois critères pour extraire les régions de couleur différentes d'une image : Ils sont définis sous une approche locale, régionale ou bien globale suivant que l'on étudie un voisinage ou une région. Ces trois critères mesurent des similarités couleur et déterminent les critères de croissance des régions. Les critères de similarité sont encore ici fixés par rapport à un seuil fixé ad hoc. Le processus de croissance de région est initialisé par le premier pixel de l'image. A chaque itération l'image est divisée en régions de couleurs similaires. Si les régions extraites ne sont pas connectées, un processus de fusion permet de regrouper les régions non contiguës mais avec une distribution colorimétrique similaire.
- Dibos [Dibos 97] combine la croissance de régions avec une méthode variationnelle (par la minimisation d'une fonctionnelle). L'algorithme cherche à

faire diminuer l'énergie globale en fusionnant éventuellement deux régions voisines. La distance utilisée est la distance de Stiles [Wyszecki 82] qui se prête le mieux à l'utilisation d'une fonctionnelle dans l'espace *RVB*.

Ces travaux portant sur la croissance de régions en segmentation d'image couleur se rejoignent sur un point : pour agréger un pixel à une région, une mesure de similarité couleur est calculée et le critère de croissance (i.e. de fusion du pixel à la région candidate) est paramétré par un seuil qui est fixé de façon heuristique. Ce seuil semble être un des seuls points faibles de ces méthodes car l'alliance de la croissance de région avec la similarité couleur (éventuellement couplée à d'autres critères) permet d'extraire de façon assez robuste les différentes régions d'une image.

✓ La Morphologie Mathématique

La ligne de partage des eaux (L.P.E) est l'outil de base pour la segmentation d'images par morphologie mathématique [Vincent 90, Serra 82, Coster 89]. Il existe différents algorithmes pour le calcul de la L.P.E [Vincent 90] qui opère généralement sur un gradient de l'image. Pour utiliser la couleur à travers la L.P.E, plusieurs choix sont possibles [Saarinen 95] : on peut calculer la L.P.E sur chacune des composantes couleur séparément et combiner ensuite les résultats. On peut calculer la L.P.E directement sur l'image couleur ou encore sur le gradient de l'image couleur. Meyer [Meyer 92] a proposé un algorithme de L.P.E pour segmenter les images couleur qui s'apparente à une méthode de croissance de régions. La croissance de régions est initialisée à partir de marqueurs obtenus par la sélection de tous les minima du gradient couleur. Mais cette obtention des marqueurs peut se faire par n'importe quelle méthode morphologique. Meyer utilise les différences de couleur entre les pixels qui sont ordonnés dans une file afin d'établir une croissance de régions rapide. Meyer calcule la L.P.E sur l'image couleur initiale mais d'autres travaux [Belhomme 97, Elmoataz 96, Salembier 96] ont modifié cette méthode et utilisent le gradient couleur ainsi que l'image couleur initiale. La mesure de similarité entre un pixel et le vecteur couleur moyen de la région est réalisé dans l'espace couleur *RVB* par une distance euclidienne. Evidemment les similarités d'un tel algorithme avec les méthodes de croissance de régions sont nombreuses et la L.P.E couleur s'affilie plus à un algorithme optimal de croissance de régions qu'à une méthode morphologique. Cette méthode a été utilisée pour segmenter diverses classes d'images [Shafarenko 97]. Le problème principal de la L.P.E réside dans la sur segmentation dus aux minima parasites (souvent causés par le bruit) que l'on peut trouver dans les images couleur. Cela peut s'éviter par l'usage de marqueurs [Vincent 90], de filtrage de l'image initiale [Elmoataz 96] ou bien par la construction du Graphe d'Adjacence des Régions [Saarinen 94] suivi de la fusion des régions suivant une mesure de similarité (taille, forme, couleur etc.).

✓ La classification de pixels

En classification de pixels, on trouve les méthodes qui sont généralement utilisées en analyse d'images. Cela consiste à affecter à chaque pixel de l'image une classe qui définit les régions à extraire de l'image.

↳ Les réseaux de neurones sont souvent utilisés et les différences entre les travaux se situent généralement aux niveau des réseaux utilisés (hopfield, rétropropagation, minimisation de fonctionnelles) et des paramètres d'entrée. On notera les travaux de [Campadelli 97, Littman 97, Gaio 94, Sammouda 97, Raghu 95, Delmotte 97]

↳ Des méthodes de classification sont employées afin d'obtenir un prétraitement pour une croissance de régions. Ceci peut se faire en utilisant le k-means [Gevers 90] ou bien le Minimum Volume Ellipsoid [Comaniciu 97].

✓ Les autres méthodes

On trouve également différentes autres méthodes :

↳ Dans les méthodes utilisant les champs de Markov aléatoires [Lui 94, Wright 89, Daily 89], une image couleur est considérée comme un ensemble de trois images monochromes couplées par l'intermédiaire de champs markoviens. Le formalisme des champs markoviens permet pour une image, d'exploiter les propriétés de continuité par morceaux de couleur en utilisant seulement des relation locales.

↳ La transformée de Karhunen-Loève est employée par Umbaugh et Guzmán de León [Umbaugh 93, Guzmán de León 96], ils utilisent cette méthode et segmentent ensuite leurs images couleur en les divisant suivant l'axe de plus grande variance par la méthode du médian. En fait une fois la transformée de Karhunen-Loève calculée sur une image couleur, on peut appliquer n'importe quelle méthode de segmentation couleur et celle-ci s'effectuera dans l'espace $X_1 X_2 X_3$ [MacAulay 89, Tomczak 98].

↳ Une autre stratégie de segmentation consiste à fusionner les informations apportées par les trois composantes couleur d'un espace couleur quelconque afin de définir une information unique traduisant soit un critère d'homogénéité soit un critère de discontinuité. On peut alors utiliser les algorithmes usuels développés pour les images en niveaux de gris. Se référer aux travaux de Carron et Smolle [Carron 94, Smolle 96].

↳ On trouve des approches utilisant des « *templates* » couleur. Celles-ci permettent d'avoir des descriptions compactes de chaque objet à reconnaître. Une mesure de similarité entre les objets et ces « *templates* » de référence permet la reconnaissance des objets décrits dans la base [Brock-Gunn 95].

2.1.2. La ligne de partage des eaux

Au vu des méthodes utilisées en analyse d'image, nous avons décidé de nous placer dans le cadre de la morphologie mathématique. Les méthodes appartenant à cette discipline nous semblent très adaptés pour segmenter des images couleur surtout lorsque celles-ci présentent une forte complexité. La ligne de partage des eaux, la principale transformation en morphologie mathématique, a été utilisée pour segmenter des images en niveaux de gris dans un premier temps. Puis son extension aux images couleur présentée par [Meyer 92] a permis

d'ouvrir une nouvelle voie dans le traitement des images couleur. Avant de considérer la ligne de partage des eaux couleur, nous allons effectuer quelques rappels de notions de morphologie mathématique et décrire la méthode de segmentation par la ligne de partage des eaux sur des images en niveaux de gris.

Pour bien comprendre le principe de la L.P.E il nous faut comparer une image en niveaux de gris à un relief terrestre, l'altitude jouant le rôle des niveaux de gris. Dans cette comparaison, une image peut se diviser en plusieurs zones. Intuitivement, une goutte d'eau tombant sur le relief de l'image descend en suivant la pente la plus prononcée jusqu'à atteindre un minimum. L'ensemble de tous les points menant au même minimum est appelé un bassin versant i.e. un bassin versant correspond à l'espace géographique dans lequel toutes les eaux convergent vers le même minimum. Par cette comparaison une image peut se diviser en plusieurs zones distinctes. Les lignes séparant les différents bassins versants sont appelées les lignes de partage des eaux. L'idée pour segmenter une image est de déterminer cette ligne de partage des eaux sur une image c'est à dire les lignes le long desquelles les niveaux de gris varient rapidement. Nous allons voir une façon de définir la ligne de partage de eaux (définition par immersion) proposée par [Vincent 90].

2.1.2.1. Définitions de base

Considérons une image I définie discrète, i.e. définie sur une trame (hexagonale ou carrée) à valeurs discrètes et dont le domaine de définition est $D_I \subset \mathbb{Z}^2$. I est donc supposée ne prendre que des valeurs discrètes dans l'intervalle $[0, N]$, N étant un entier positif arbitraire.

$$I \left(\begin{array}{l} D_I \subset \mathbb{Z}^2 \rightarrow \{0, 1, \dots, N\} \\ p \rightarrow I(p) \end{array} \right)$$

Ce type d'image va nous permettre de définir les bassins versants associés à chaque minimum régional de manière simple. Nous suivons ici la présentation de [Vincent 90] et noterons G la trame qui peut être de n'importe quel type : une trame carrée en quatre ou huit voisinage, ou bien une trame hexagonale en six voisinage. G est un sous-ensemble de $\mathbb{Z}^2 \times \mathbb{Z}^2$.

Définition 1 : Un chemin P de longueur l entre deux pixels p et q dans une image I est un ensemble de $(l+1)$ pixels $(p_0, p_1, \dots, p_{l-1}, p_l)$ tel que $p_0 = p, p_l = q$, et $\forall i \in [1, l], (p_{i-1}, p_i) \in G$.

Par la suite on notera $l(P)$ la longueur d'un chemin donné P ainsi que $N_G(p)$ l'ensemble des voisins du pixel p suivant la trame G : $N_G(p) = \{p' \in \mathbb{Z}^2, (p, p') \in G\}$.

Avant d'introduire la notion de ligne de partage des eaux rappelons la notion de minimum.

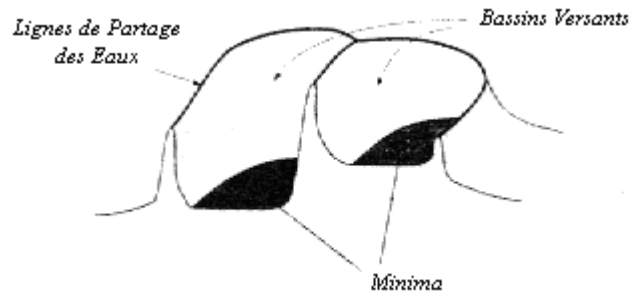


Figure 13. Minima, bassins versants et ligne de partage des eaux [Vincent 90].

Définition 2 : Un minimum M de I d'altitude h est un plateau de pixels connectés de valeur h à partir duquel il est impossible d'atteindre un point de moindre altitude sans avoir à remonter :

$$\begin{aligned} \forall p \in M, \quad \forall q \notin M, \quad \text{tel que } I(q) \leq I(p) \\ \forall P = (p_0, p_1, \dots, p_l) \text{ tel que } p_0 = p \text{ et } p_l = q \\ \exists i \in [1, l] \text{ tel que } I(p_i) > I(p_0). \end{aligned}$$

A tout minimum sera donc associé la zone géographique d'où une goutte d'eau, suivant la ligne de plus grande pente, arrivera dans ce minimum. Ceci définit alors les bassins versants qui sont les zones d'influence d'un minimum (Figure 13).

Définition 3 : Le bassin versant associé à un minimum M est l'ensemble des pixels p de D_I tel qu'une goutte d'eau tombant au point p suit le relief suivant la ligne de plus grande pente pour arriver en M .

Les lignes séparant les différents bassins versants est appelée ligne de partage des eaux.

2.1.2.2. Définition par immersion

Pour construire la ligne de partage des eaux, nous allons procéder en immergeant progressivement dans l'eau l'image I par le bas. Pour que l'eau progresse de façon régulière sur toute l'image, supposons que chaque minimum régional soit « percé ». Nous allons ensuite immerger progressivement l'image jusqu'à son immersion complète. En partant des minima de plus faible altitude, l'eau va progressivement remplir les bassins versants de I . En partant des ouvertures que nous avons créées l'eau va progresser dans chacun des bassins versants correspondant à chaque minimum. Puis lorsque deux eaux en provenance de deux minima différents se rencontrent, on construit un barrage, afin que des eaux d'origines différentes ne se mélangent pas (Figure 14).

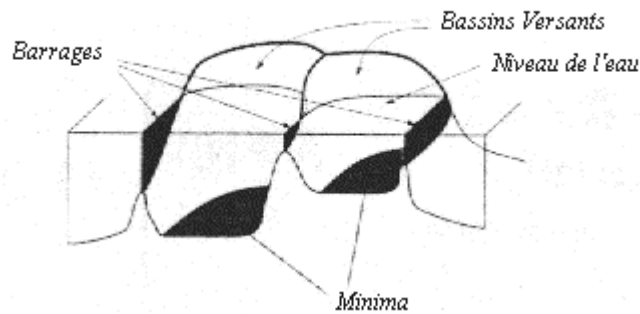


Figure 14. Barrages aux endroits où l'eau provenant de deux minima se mélangeraient [Vincent 90].

A la fin de la procédure d'immersion, chaque minimum est alors complètement entouré de barrages qui délimitent le bassin versant qui lui est associé (i.e. la zone d'influence de ce minimum). L'ensemble des barrages constitue la ligne de partage des eaux de l'image I . Exprimons à présent ce processus de façon plus formelle : Pour une image I , soient h_{min} et h_{max} respectivement la plus petite et la plus grande valeur prise par I sur son domaine D_I . Notons $T_h(I)$ le seuillage de I à l'altitude h :

$$T_h(I) = \{p \in D_I, I(p) \leq h\}.$$

Notons $C(M)$ le bassin versant associé au minimum M et $C_h(M)$ le sous-ensemble de ce bassin versant constitué des points d'altitude plus petite ou égale à h :

$$C_h(M) = \{p \in C(M), I(p) \leq h\} = C(M) \cap T_h(I).$$

On notera $min_h(I)$ l'ensemble des points étant des minima d'altitude h . Avant de définir la ligne de partage des eaux, nous devons rappeler les définitions de la distance géodésique et de la zone d'influence géodésique.

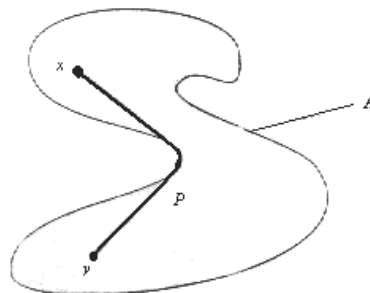


Figure 15. La distance géodésique entre x et y à l'intérieur de A est le plus petit chemin entre ces deux points qui est totalement inclus dans A [Vincent 90].

Définition 4 : La distance géodésique $d_A(x, y)$ entre deux pixels x et y appartenant à un objet A , est le plus petit arc d'extrémités x et y qui soit totalement inclus dans A (Figure 15):

$$d_A(x, y) = \inf \{l(P) = (p_0, \dots, p_l) / p_0 = x \text{ et } p_l = y, \forall p_i \in P \text{ avec } i \in [1, l-1], p_i \in A\}$$

Supposons maintenant que A contienne un ensemble B composé de plusieurs composantes connectées B_1, B_2, \dots, B_k .

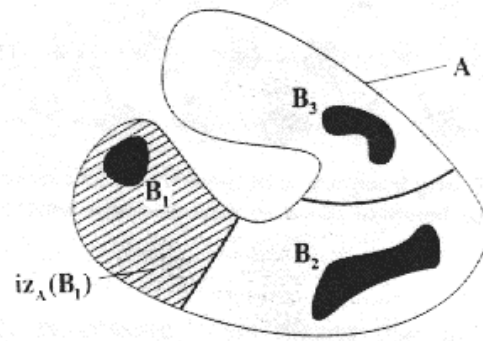


Figure 16. La zone d'influence géodésique de composantes connectées B_i à l'intérieur d'un ensemble A [Vincent 90].

Définition 5 : La zone d'influence géodésique $iz_A(B_i)$ d'une composante B_i de B dans A est la zone où les points de A ont leur distance géodésique à l'ensemble B_i plus petite que leur distance géodésique à n'importe quel autre composante de B (Figure 16):

$$iz_A(B) = \left\{ p \in A, \forall j \in [1, k] / \{i\}, d_A(p, B_i) < d_A(p, B_j) \right\}.$$

$$IZ_A(B) = \bigcup_{i \in [1, k]} iz_A(B_i)$$

Nous pouvons maintenant définir la ligne de partage des eaux.

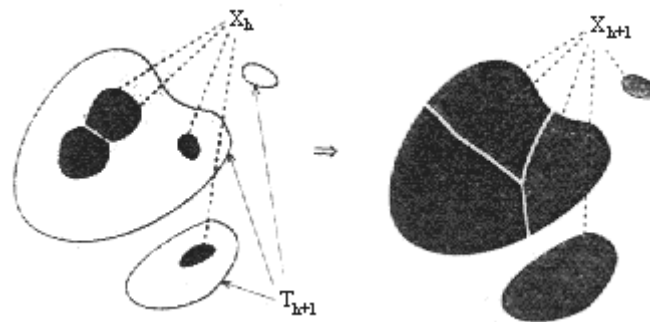


Figure 17. Itération entre X_h et X_{h+1} [Vincent 90].

Définition 6 : L'ensemble des bassins versants de l'image I est égal à l'ensemble $X_{h_{\max}}$ obtenu après l'itération suivante :

- 1) $X_{h_{\min}} = T_{h_{\min}}(I)$
- 2) $\forall h \in [h_{\min}, h_{\max} - 1], X_{h+1} = \min_{h+1}(I) \cup IZ_{T_{h+1}(I)}(X_h).$

La ligne de partage des eaux correspond au complémentaire de cet ensemble dans D_I , i.e. à l'ensemble des points de D_I qui n'appartiennent à aucun bassin versant. Le passage entre deux itérations est illustré par la Figure 17.

2.1.2.3. La L.P.E contrainte par marqueurs

La L.P.E calculée directement sur une image réelle présente une sur-segmentation importante, due au bruit qui crée de nombreux minima parasites. Il est donc nécessaire de mettre en place des procédures permettant de les éliminer. Supposons que nous connaissions un ensemble connexe de points par objets à isoler ainsi qu'un ensemble de points de l'extérieur. On appellera ces composantes connexes des marqueurs. Si l'on parvenait à modifier l'image sur laquelle on calcule la L.P.E en lui imposant ces marqueurs comme minima régionaux, on obtiendrait une L.P.E présentant une boucle autour de chacun des objets, chaque bassin versant correspondant soit à un unique objet soit au fond de l'image. Cette technique de L.P.E contrainte par marqueurs permet de s'affranchir de la recherche exacte des contours et garantit le nombre de contours que l'on trouvera : un autour de chaque composante connexe. Notons que nous n'avons pas précisé comment obtenir ces marqueurs et c'est la partie la plus délicate. La difficulté est reportée sur la détermination des marqueurs qui permettent de localiser grossièrement les objets. Pour résumer la L.P.E contrainte par marqueurs fonctionne de la manière suivante :

1. Trouver les marqueurs, i.e. une composante connexe par objet et une composante connexe pour le fond de l'image,
2. Déterminer l'image sur laquelle on calcule la L.P.E,
3. Imposer les minima,
4. Calculer la L.P.E.

2.1.3. La ligne de partage des eaux couleur

L'extension de la ligne de partage des eaux en niveaux de gris à la couleur est relativement récente. L'innovation principale vint des travaux de [Meyer 92] qui proposa un algorithme dont les similarités avec les méthodes de croissance de régions sont nombreuses. L'algorithme proposé par [Meyer 92] consiste en une croissance de régions. Il utilise des différences de couleur entre pixels et une structure de données de type file hiérarchique pour réaliser une croissance de régions rapide où les régions se propagent à partir de marqueurs. Ce procédé calcule la ligne de partage des eaux directement sur l'image couleur d'origine. On dispose alors d'un algorithme de croissance de régions en ce sens qu'il est initialisé par des marqueurs qui identifient l'intérieur des régions et ces régions sont étendues jusqu'à occuper tout l'espace disponible. Dans l'algorithme la croissance de régions consiste à extraire un pixel de la file de priorité et à examiner ses régions voisines. Une distance entre le pixel et les régions voisines est alors calculée et le pixel est fusionné avec la région la plus proche (i.e. la plus similaire). On note ici une similarité avec les méthodes de croissance de régions à travers cette mesure de distance. Celle-ci est justement un paramètre important pour estimer la probabilité d'appartenance d'un pixel à une région. C'est la modification de cette fonction distance qui a entraîné des recherches actives. [Salembier 96] a modifié cette fonction afin d'améliorer la précision de l'algorithme et calcule la L.P.E directement sur le gradient de l'image couleur. D'autres méthodes [Saarinen 94, Elmoataz 96, Belhomme 97] ont modifié cette fonction distance en y ajoutant d'autres mesures liées à des critères locaux et globaux. Ceci a été modélisé par [Belhomme 97] à travers la définition formelle suivante. Une mesure

de similarité entre un pixel et une région voisine s'exprime alors de la manière suivante par la fonction d'agrégation :

$$\phi(p, R) = \left(\sum_{i=1}^l \alpha_i \cdot \phi_i(p) + \sum_{i=l+1}^m \alpha_i \cdot \phi_i(p, R) \right)$$

où les $\phi_i(p)$ représentent l critères locaux directement extraits d'une image (image couleur, gradient, ...) et les $\phi_i(p, R)$ représentent $m-l$ critères globaux résultant de comparaisons statistiques entre un point p et une de ses régions voisines R . Les α_i sont des coefficients de pondération qui respectent la propriété suivante : $\sum_{i=1}^m \alpha_i = 1$.

On retrouve la définition de la ligne de partage des eaux en prenant $l = m = 1$ et $\phi_i(p) = |\text{grad}(I)|$.

2.2. Notre approche de la segmentation couleur

Nous avons vu que la ligne de partage des eaux est un outil puissant pour segmenter des images couleur, et nous avons choisi de baser nos travaux sur cette méthode. Nous présentons dans un premier temps notre fonction d'agrégation couleur et nous nous attardons sur l'optimisation de la fonction d'agrégation grâce à la détermination de critères optimaux. Finalement, nous présentons deux expérimentations de notre approche de la segmentation couleur.

2.2.1. Notre fonction d'agrégation couleur

Au vu d'une image couleur, quelles informations nous permettent d'interpréter les informations sémantiques contenues dans cette image ? En fait ceci peut se résumer à deux informations : la couleur des objets présents sur l'image et les transitions de couleur entre les objets. Une ligne de partage des eaux utilisant ces deux informations (que l'on retrouvera toujours quelles que soient les images rencontrées) peut permettre d'extraire de façon très précise et très fiable les objets présents dans l'image. La couleur des objets présents sur l'image se traduit directement par une mesure statistique de la couleur moyenne des régions et les transitions entre les couleurs s'expriment facilement par le gradient couleur. L'utilisation de la couleur peut alors amener à des résultats qui correspondent à des changements plus significatifs dans les images couleur. Une ligne de partage des eaux couleur sera précise en utilisant les propriétés suivantes : l'information locale donnée par le gradient couleur traduisant les transitions entre les couleurs et l'information globale donnée par une mesure de la couleur moyenne des régions traduisant l'homogénéité globale des régions. Nous proposons d'exprimer ces deux critères à travers une fonction d'agrégation. Ceci se traduit par la formule suivante :

$$\phi(p, R) = (1 - \alpha) \left\| \overline{I_{C_1 C_2 C_3}}(R) - I_{C_1 C_2 C_3}(p) \right\| + \alpha \left\| \nabla I_{C_1 C_2 C_3}(p) \right\| \quad (1)$$

avec $\overline{I_{C_1 C_2 C_3}}(R)$ le vecteur donnant la couleur moyenne de la région R pour l'image I dans l'espace $C_1 C_2 C_3$, $I_{C_1 C_2 C_3}(p)$ le vecteur donnant la couleur du point p dans l'image I et $\nabla I_{C_1 C_2 C_3}(p)$ le gradient couleur au point p .

Cette fonction d'agrégation combine l'information locale (le module du gradient couleur) et l'information globale (résultant d'une comparaison statistique entre la couleur d'un

point p et une région voisine R). α est un coefficient de pondération qui permet de modifier la relation et l'influence des critères global et local durant le processus de croissance. Avec $\alpha = 1$ on retrouve l'algorithme classique de la L.P.E mais dont le gradient est calculé sur une image couleur. Avec $\alpha = 0$ on retrouve l'algorithme de [Meyer 92] qui n'utilise que la distance couleur entre un pixel et une région.

2.2.2. Optimisation de la fonction d'agrégation

Maintenant que nous avons défini notre méthode de segmentation d'images couleur en définissant une nouvelle fonction d'agrégation de pixels, définissons quels sont les points importants de la mise en place d'une telle méthode. Nous avons défini la mesure de la valeur d'une fonction d'agrégation entre les pixels par (1) mais sans toutefois préciser quel espace couleur choisir, ni comment calculer le gradient couleur dans cet espace. Il nous reste également à définir la comparaison statistique couleur entre un pixel et une région. Enfin α est un paramètre déterminant dans l'utilisation de notre méthode et il reste à définir. Enfin nous n'avons pas précisé comment obtenir les marqueurs. Pour conclure l'utilisation de notre méthode nécessite plusieurs choix :

- Choix de l'espace couleur,
- Choix d'une métrique pour comparer $\overline{I_{C_1C_2C_3}}(R)$ et $I_{C_1C_2C_3}(p)$,
- Choix d'une méthode de calcul du gradient,
- Choix des marqueurs,
- Choix du coefficient de pondération α .

Afin de bien réaliser ces choix qui sont déterminants pour obtenir de bons résultats, nous proposons de définir une stratégie d'optimisation. Elle se déroule en plusieurs étapes successives, chacune amenant à l'établissement d'un paramètre ou des données d'entrée de notre méthode. Chacune des étapes utilise soit des mesures quantitatives ou bien des informations a priori sur les images couleur. C'est pourquoi le problème doit être complètement connu. C'est à dire que tous les cas pouvant être rencontrés doivent représenter des configurations couleur différentes dans les images. Pour bien définir le problème et il faut étudier préalablement son contexte. Une fois le problème complètement déterminé la stratégie peut être appliquée :

1. Choix d'une métrique pour comparer $\overline{I_{C_1C_2C_3}}(R)$ et $I_{C_1C_2C_3}(p)$.

Ici le choix dépend d'informations a priori, pour des objets homogènes on utilisera une distance euclidienne mais la distance de Mahalanobis est mieux adaptée pour des régions qui présentent de fines variations étant donné que la différence calculée entre la couleur d'un pixel et celle d'une région R est rapportée à l'écart type de la couleur de cette dernière. La mesure de distance que nous avons utilisée est une mesure euclidienne. Pour mesurer la distance entre deux couleurs A et B dans un espace couleur $C_1C_2C_3$ on utilisera la formule suivante :

$$d(A, B) = \sqrt{(A_{C_1} - B_{C_1})^2 + (A_{C_2} - B_{C_2})^2 + (A_{C_3} - B_{C_3})^2} .$$

A_{C_i} étant la valeur de la composante numéro i de la couleur A .

2. Extraction des marqueurs.

La couleur donnant une information a priori sur les objets à isoler dans une image, l'extraction des marqueurs doit se faire sur des composantes qui correspondent à

l'information recherchée. En disposant d'une connaissance a priori sur l'information recherchée dans une image, nous allons utiliser tel espace couleur. Par exemple si une scène présente une propriété directement liée à une composante d'un espace couleur, on préférera cette composante. En fait l'extraction des marqueurs doit utiliser de façon combinée différentes composantes de différents espaces couleur et peut se faire sur une seule composante couleur où sur la combinaison de plusieurs composantes de différents espaces couleur. Cette étape est très importante étant donné que la principale difficulté dans une méthode de croissance de régions réside dans l'extraction des marqueurs et la couleur améliore de façon significative cette extraction. Pour le choix des composantes couleur à combiner, l'utilisation d'informations a priori sur l'information recherchée est indispensable. Nous verrons par la suite sur différents exemples comment procéder.

3. Choix du coefficient de pondération α

Ce choix se fait également vis à vis des connaissances a priori sur les objets à extraire.

- Si les contours des objets sont très nets on prend α avec une valeur entre 0,5 et 1,
- Si les contours sont moins nets et les régions peu homogènes en couleur on prend α proche de 0,5,
- Si les contours ne sont pas nets mais que les régions sont homogènes, on prend α proche de 0.

Il nous reste finalement deux critères à optimiser qui sont l'espace couleur et le gradient couleur. Nous allons maintenant voir en détail comment réaliser ces choix.

2.2.3. Choix de l'espace couleur

Comment définir quel espace couleur utiliser ? Comment déterminer quel espace couleur est le mieux adapté au problème posé ? Ceci nous paraît être la principale difficulté lorsque l'on parle de choix d'un espace couleur : ce choix doit-il se faire indépendamment du problème à résoudre ou doit-il tenir compte des objets à extraire dans l'image ? Une première approche nous pousse à trouver le meilleur espace de représentation d'une image et ensuite à choisir la méthode à utiliser pour segmenter l'image, il paraît toutefois difficile de séparer le choix de l'espace couleur du problème à résoudre. Nous avons donc axé notre réflexion en ce sens : le choix de l'espace couleur est lié à la méthode de segmentation. La fonction d'agrégation que nous avons définie utilise implicitement la notion de distance à travers le gradient et à travers la métrique utilisée pour comparer $\overline{I_{C_1C_2C_3}}(R)$ et $I_{C_1C_2C_3}(p)$. Nous allons étudier les espaces couleur selon leur influence sur la notion de distance puis nous présenterons une méthode de choix d'un espace couleur utilisant une segmentation manuelle des objets comme information a priori. Enfin nous proposerons une approche permettant de définir une représentation optimale de la couleur par apprentissage à partir d'informations a priori.

2.2.3.1. Comparaison des distances selon l'espace couleur

La notion de distance entre les couleurs se traduit généralement par une distance euclidienne. Voici les distances pour les espaces couleur que nous avons vus précédemment [CIE 76, Chassery 84, Schettini 93, Lasserre 97, Tremeau 97, Wyszecki 82] :

$$\begin{aligned}
 RVB : \Delta d &= \sqrt{\Delta R^2 + \Delta V^2 + \Delta B^2} \\
 XYZ : \Delta d &= \sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2} \\
 L^* u^* v^* : \Delta d &= \sqrt{\Delta L^{*2} + \Delta u^{*2} + \Delta v^{*2}} \\
 L^* a^* b^* : \Delta d &= \sqrt{\Delta L^{*2} + \Delta a^{*2} + \Delta b^{*2}} \\
 HSL : \Delta d &= \sqrt{\Delta H^2 + \Delta S^2 + \Delta L^2} \\
 Chassery : \Delta d &= \sqrt{\Delta A^2 + \Delta s^2 + \left(\frac{s_1 + s_2}{2}\right)^2 \Delta t^2}
 \end{aligned}$$

Cette distance est évidemment liée aux performances de l'espace couleur. Elle est meilleure si l'espace est perceptuellement uniforme, à savoir qu'une petite variation dans une des composantes est perceptible de façon approximativement égale sur l'échelle des valeurs (deux couleurs perceptuellement proches doivent l'être au niveau de leurs composantes dans l'espace associé) [Kunt 91].

Si un espace de couleur est uniforme, les différentes teintes d'une même couleur doivent être linéairement distantes d'une couleur donnée. On saisit alors l'importance de la linéarité ou non des équations définissant les espaces couleur. Pour évaluer la performance de la distance dans chacun des espaces couleur que nous avons détaillés, nous examinons pour une image donnée la distance de chaque couleur de l'image au blanc de référence de chaque espace :

Soit p , de coordonnées (x_p, y_p) , un point d'une image couleur I .

Calculer $\Delta d = \text{dist}_{C_1 C_2 C_3}(p, B)$ où B est le point représentant le blanc dans l'espace $C_1 C_2 C_3 \in \{RVB, XYZ, L^* u^* v^*, L^* a^* b^*, HSL, Chassery\}$

Sur l'image originale nous avons rajouté une zone totalement blanche (tous les points de cette zone ont pour couleur $(255,255,255)$ en RVB) sur une zone paraissant blanche (de teinte équivalente). Ceci va nous permettre de vérifier également la distance entre le blanc et une gamme de couleur blanche contenue dans l'image.

Sur les images résultantes, qui sont donc des images de distance, une zone claire représente une zone proche du blanc dans l'espace couleur et une région foncée représente une région éloignée du blanc. Pour qu'un espace couleur soit uniforme, on doit avoir, pour des zones de couleur proches, approximativement les mêmes niveaux de gris. L'image sur laquelle nous avons réalisé cette comparaison est une image de serviettes de différentes couleurs (blanche, rouge, jaune et bleue) sur une table de fond marron.

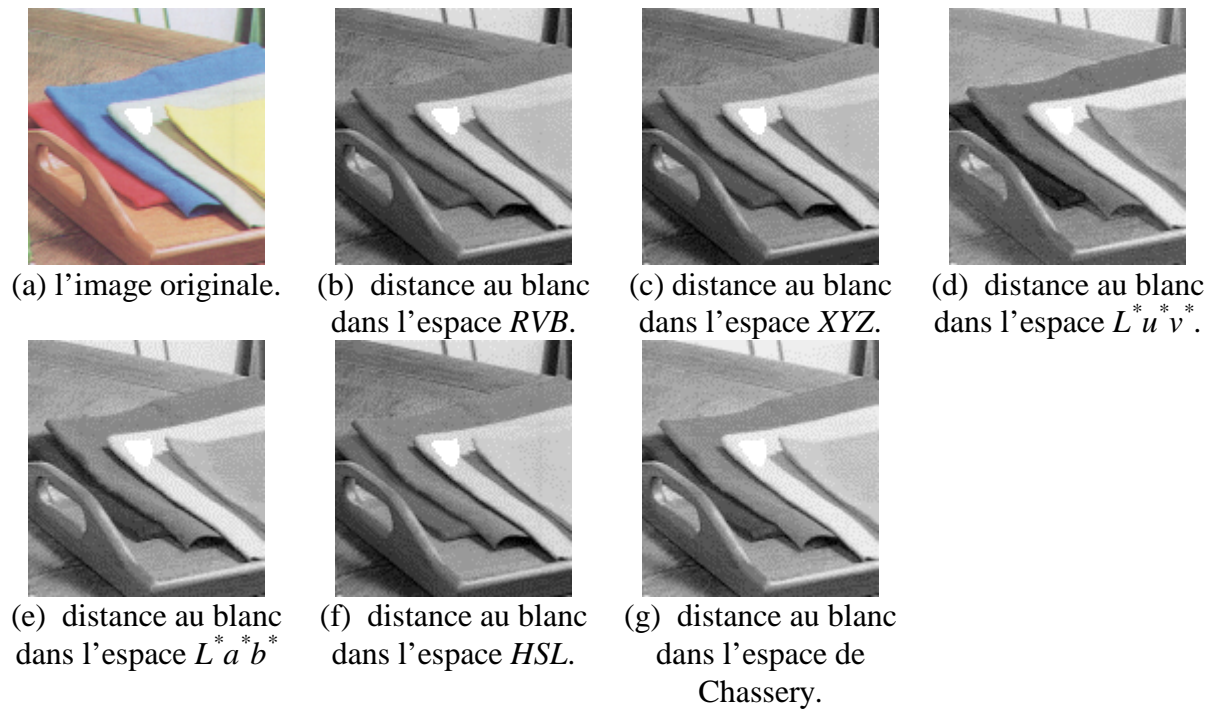


Figure 18. L'image originale et les images de distance au blanc dans chaque espace couleur.

Afin de visualiser l'influence de l'espace couleur sur la mesure de la distance entre les couleurs, nous avons isolé sur les histogrammes de distance des Figures 18 (b) à (g) les différents régions présentes sur l'image originale, à savoir : les serviettes (bleue, rouge, jaune, blanche) ainsi que les zones marron et les ombres (Figure 19).

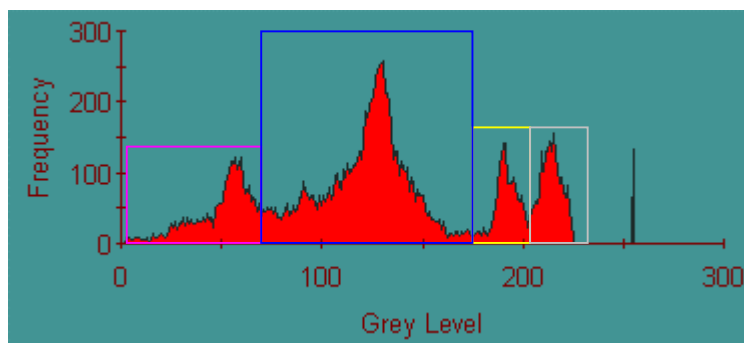


Figure 19(a) : Histogramme de l'image de distance 18(b) (espace RVB). La zone marron est contenue dans les zones bleues et rouges.

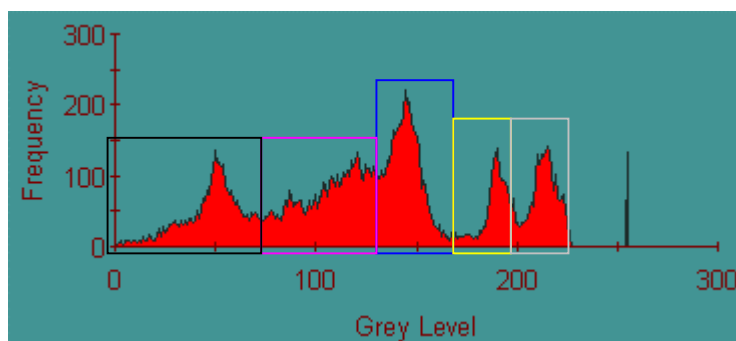


Figure 19(b) : Histogramme de l'image de distance 18(c) (espace XYZ).

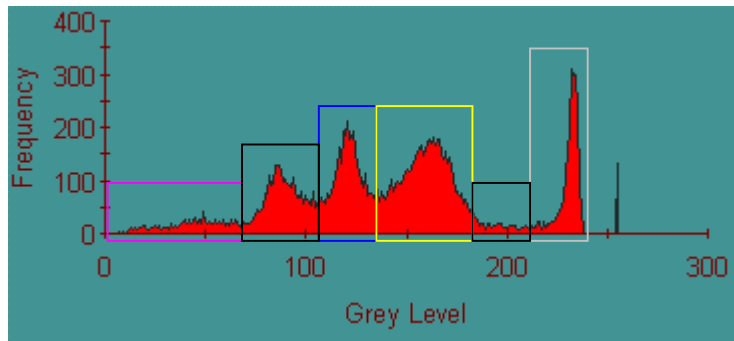


Figure 19(c) : Histogramme de l'image de distance 18(d) (Espace $L^*u^*v^*$).

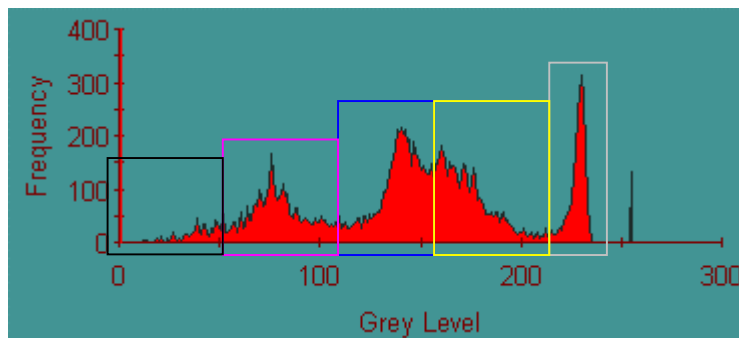


Figure 19(d) : Histogramme de l'image de distance 18(e) (Espace $L^*a^*b^*$). La zone marron est confondue avec les zones bleues et jaunes.

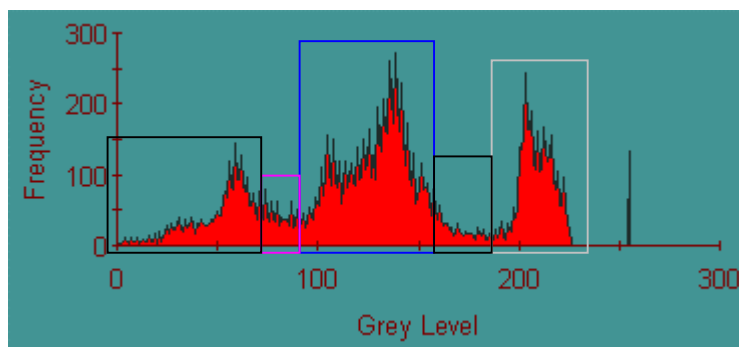


Figure 19(e) : Histogramme de l'image de distance 18(f) (Espace HSL). Les zones blanches et jaunes sont confondues.

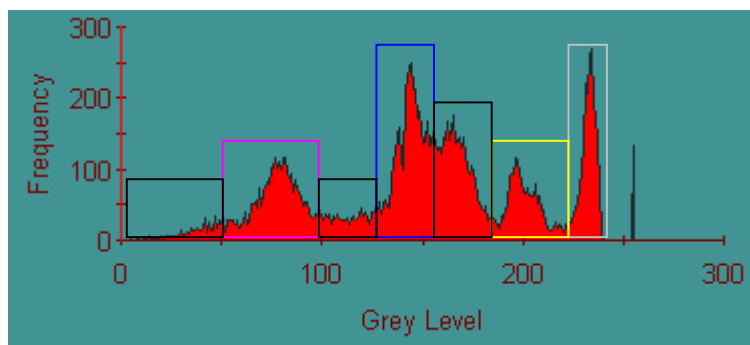


Figure 19(f) : Histogramme de l'image de distance 18(g) (Espace de Chassery).

Figure 19. histogrammes des Figures 18 (b) à (g).

Les zones de l'image sont données dans les histogrammes par les couleurs suivantes : serviettes bleue (—), jaune (—), rouge (—), blanche (—), les zones marron, ombres et table (—).

Grâce à la visualisation de la position et de la localisation des régions dans les histogrammes, cela nous permet d'estimer la « précision » de la métrique dans un espace couleur à savoir si une variation couleur réelle et visible se traduit directement en terme de distance. Sur les images le pic le plus à droite désigne la zone blanche présente sur l'image.

Au vu des images et des histogrammes on peut grouper ici les espaces couleur en trois groupes : *RVB* et *XYZ*, *HSL* et Chassery et enfin $L^*u^*v^*$ et $L^*a^*b^*$.

- Les systèmes *RVB* et *XYZ* se comportent sensiblement de la même manière. Les zones blanches, jaunes, bleues et rouges sont bien déterminées mais les zones de marron ne sont pas assez bien discriminées par rapport aux autres couleurs. On remarque donc que pour les autres zones homogènes de couleur qui présentent des variations de teinte et de saturation (le marron par exemple), ces deux systèmes montrent leurs limites. Le problème de ces systèmes est que chacun des axes chromatiques portent en eux une information de luminance.
- L'espace *HSL* se comporte sensiblement de la même façon que les espaces *RVB* et *XYZ* mais il est meilleur dans les zones de couleur homogène tout en étant moins performant que les autres espaces ($L^*u^*v^*$, $L^*a^*b^*$, Chassery) dans ces mêmes zones. L'espace *HSL* rend mieux compte des différences de couleur dans des zones où la teinte varie beaucoup. Cet espace est meilleur que les espaces traditionnels mais n'est pas encore assez précis dans les zones de couleur homogène. L'espace de Chassery est encore meilleur que l'espace *HSL*. Il est plus fin dans les zones de couleur homogène (notamment au niveau du blanc et du marron) mais il accentue moins les zones d'ombres (le noir est donc moins éloigné du blanc). Ce système réagit mieux que le système *HSL* au niveau de l'homogénéité des couleurs, sa dispersion spatio-colorimétrique est plus fine et tend à le rapprocher d'espaces uniformes tels que $L^*u^*v^*$ et $L^*a^*b^*$.
- Les systèmes $L^*u^*v^*$ et $L^*a^*b^*$ réagissent très bien dans les zones de couleur homogènes même si elles présentent des variations de teinte ou de saturation. Chaque pic de l'histogramme correspond à une région de couleur uniforme de l'image. On notera quelques différences entre $L^*u^*v^*$ et $L^*a^*b^*$ notamment pour le rouge qui est plus éloigné du blanc pour ce premier et il discerne de façon plus nuancée les zones de marron.

Cette discussion nous permet de nous rendre compte que suivant l'espace couleur utilisé la distance entre les couleurs ne traduit pas à fortiori une différence nette entre des couleurs pourtant visiblement éloignées.

2.2.3.2. Un critère de choix de l'espace couleur basé sur le contraste

Nous désirons définir une méthode systématique qui se base sur des mesures quantitatives pour classer des espaces couleur selon leur pertinence pour un problème donné. Pour cela nous allons utiliser deux informations. La première est une information a priori : on va isoler de façon manuelle les objets à segmenter dans l'image. Ensuite pour caractériser ou plutôt pour mesurer le degré de pertinence d'un espace par rapport à un problème donné, nous allons utiliser la distance entre les couleurs qui caractérise la dispersion des couleurs dans une image. Ceci permettra de caractériser les transitions entre les couleurs.

La méthode que nous avons définie est basée sur l'extraction du contraste d'une image. La première étape consiste à segmenter manuellement l'image afin d'isoler les zones ou objets à reconnaître. Ensuite nous calculons le contraste couleur local de l'image dans chacun des espaces couleur $C_1C_2C_3$ parmi RVB , XYZ , $L^*u^*v^*$, $L^*a^*b^*$, HSL , Chassery puis nous calculons le contraste global (somme des contrastes locaux) pour les points se trouvant aux frontières des régions extraites à l'étape précédente. Les contrastes couleur local et global sont obtenus de la manière suivante :

Soit $P(x,y) \in I$ avec I une image couleur (\forall son espace couleur).

$$c(P) = \left\{ \max(\text{dist}(P, M)) \mid M \text{ soit un } 8\text{-voisin de } P \right\}$$

Le contraste couleur global est alors : $C = \sum(c(P))$ tel que P soit un point de frontière entre deux régions.

Pour chaque espace couleur, les images couleur sont normalisées entre 0 et 255 avant le calcul des contrastes afin d'avoir des résultats équivalents dans tous les espaces couleur. La distance utilisée ici est la distance euclidienne mais n'importe quelle distance peut-être utilisée. Par contre étant donné que la détermination du meilleur espace couleur pour résoudre le problème est réalisée à l'aide de cette distance, celle-ci devra également être utilisée dans la méthode de segmentation à chaque fois que la notion de distance entre des couleurs interviendra. L'image $I_{C_1C_2C_3}$ donnant le meilleur contraste global sera celle la mieux adaptée à un traitement c'est-à-dire que sa dispersion spatio-colorimétrique correspond le mieux au problème. En effet le contraste local est fort aux endroits de transitions de couleur : à ces endroits la distance entre deux points et donc deux couleurs est plus élevée que dans des régions de couleurs homogènes. Donc si le contraste local est effectivement élevé au niveau de ces transitions, l'espace couleur est adapté au problème.

Nous avons appliqué notre méthode sur quatre images différentes à plusieurs niveaux : contraste, éclairage et contexte. Ces images sont une image d'intérieur, une image d'extérieur, une image microscopique (Figure 20) et une image constituée d'une zone très nuancée de l'image d'extérieur. Nous déduisons ensuite quel espace couleur est le plus représentatif de l'information couleur que nous avons manuellement partitionnée (Figure 21).

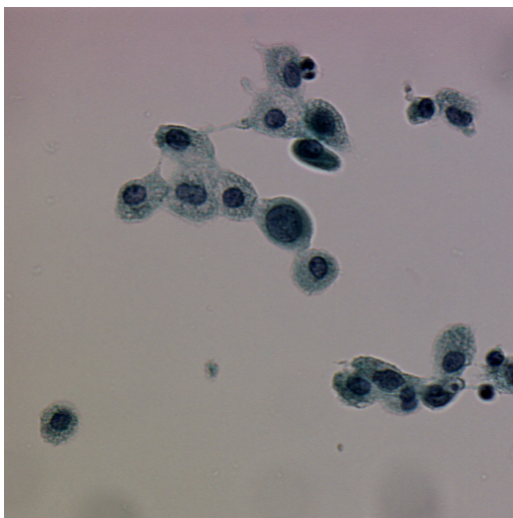
Les résultats de comparaison du contraste sont résumés dans les tableaux 2 et 3. Pour chaque espace nous avons la valeur du contraste global pour chaque image. Les différentes valeurs nous permettent de comparer les espaces.

Contexte	RVB	XYZ	$L^*u^*v^*$	$L^*a^*b^*$	HSL	Chassery	Nbre de Pts
Intérieur	171503	162207	152503	149175	144616	142757	1890
Extérieur	261714	262054	220052	209610	261340	164468	3013
Microscopique	555730	531356	373717	230138	533308	202867	8269
Ciel	55004	56100	74212	63307	48987	40439	933

Tableau 2. Tableau des contrastes globaux et nombre de points de contours par image.

<i>Contexte</i>	<i>RVB</i>	<i>XYZ</i>	<i>L*u*v*</i>	<i>L*a*b*</i>	<i>HSL</i>	<i>Chassery</i>
Intérieur	90,74	85,82	80,69	78,93	76,51	75,53
Extérieur	86,86	86,97	73,03	69,56	86,73	54,58
Microscopique	67,20	64,25	45,19	27,83	64,49	24,53
Ciel	58,95	60,13	79,54	67,58	52,50	43,34

Tableau 3. Tableau des contrastes globaux rapportés au nombre de points de contour par image.



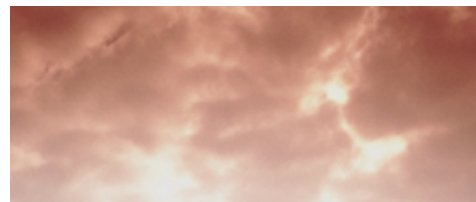
(a) contexte microscopique.



(b) contexte extérieur.



(c) contexte intérieur.



(d) contexte extérieur.

Figure 20. Les images provenant de contextes microscopique, extérieur et intérieur.

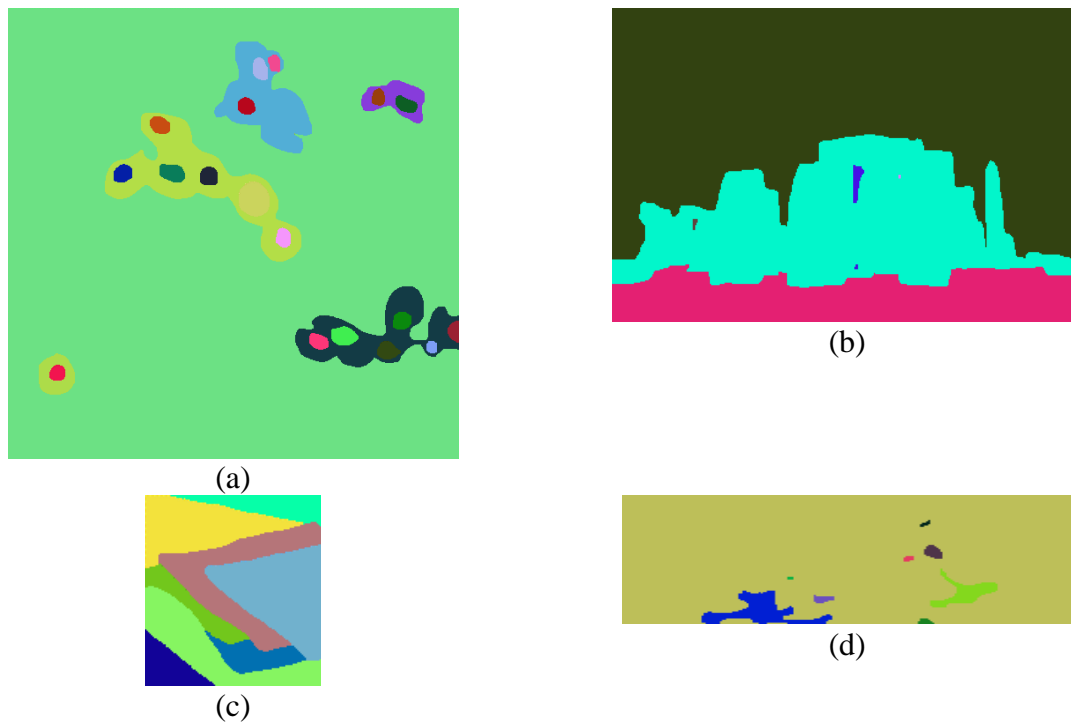


Figure 21. Les régions des images des Figure 20(a), (b) (c) et (d).

Par notre méthode l'espace couleur le mieux adapté au problème est celui qui donne la plus grande valeur de contraste. Cette aide au choix de l'espace couleur peut permettre un choix systématique basé sur des données quantitatives pour résoudre un problème. Une fois l'espace optimal déterminé, seule la méthode de segmentation peut varier. On remarque que suivant les images, ce ne sont pas toujours les mêmes espaces couleur qui donnent les meilleurs résultats. On peut penser cependant que même si les transitions couleur sont moins prononcées en terme de distance entre les couleurs dans certains espaces, elles donnent des résultats acceptables (ce qui a été vérifié par les travaux détaillés d'Ohta sur la segmentation couleur [Ohta 80]). Tout dépend donc du niveau de précision que l'on veut atteindre. Sur nos images, des espaces uniformes tels que $L^*u^*v^*$ ou $L^*a^*b^*$ ne deviennent « utiles » que lorsque les différences entre les couleurs sont difficiles à percevoir. Dans les autres cas des espaces non uniformes peuvent suffire.

Pour choisir l'espace couleur, on constituera une base d'images représentatives du problème de segmentation. Ces images sont segmentées manuellement et l'on calcule le contraste couleur pour chacune d'entre elles dans les différents espaces couleur. On calcule ensuite la moyenne des contrastes dans un même espace sur les images de la base, l'espace couleur choisi est celui qui donne le contraste maximal.

On notera que les images choisies ne reflètent pas forcément un problème réel de segmentation d'images. Le but ici n'est pas de segmenter totalement une image c'est à dire d'extraire tous les objets présents sur une image : chacune des images a été segmentée manuellement en utilisant des informations a priori sur les objets. Notre méthode de choix de l'espace couleur est adaptée à notre méthode de segmentation et dépend des objets de référence choisis. C'est ce dernier point qui importe : le choix de l'espace couleur dépend du problème et par conséquent des objets à extraire. Ce choix ne peut se faire en totale omission des informations a priori contenues dans les images et ne doit pas se détacher totalement d'un cadre applicatif. Pour s'en convaincre, il suffit de constater que pour l'image d'extérieur l'espace XYZ donne le meilleur contraste alors qu'en ne considérant qu'une partie de l'image

ainsi que la segmentation de référence, l'espace $L^*u^*v^*$ devient plus approprié. Les régions à extraire demandant une précision plus grande, un espace plus précis devient nécessaire.

2.2.3.3. Une représentation optimale de la couleur par apprentissage d'informations a priori

Nous avons vu une méthode de choix de l'espace couleur basée sur une mesure du contraste, une autre méthode consiste à définir une nouvelle représentation de l'information couleur. Une stratégie peut consister à utiliser la fusion de composantes couleur provenant de plusieurs espaces couleur. La principale difficulté repose sur le choix des composantes couleur à fusionner et sur l'élaboration d'un critère d'évaluation de la fusion des composantes couleur. Nous proposons d'utiliser des informations a priori sur les images couleur afin de sélectionner les composantes couleur à fusionner. L'image monochromatique résultante donne une meilleure représentation de la couleur par rapport aux informations a priori.

L'extraction d'informations a priori se fait comme au paragraphe précédent en procédant à une segmentation manuelle d'une base d'images représentatives qui décrivent le problème de segmentation.

2.2.3.3.1. Fusion de composantes couleur

En utilisant une combinaison linéaire, nous définissons la fusion de deux composantes couleur. Ceci s'exprime par la formule suivante :

$$I_{CL}(x, y) = I_{C_1C_2C_3}^{C_i}(x, y) * (1 - \alpha) + I_{C_1C_2C_3}^{C_i'}(x, y) * \alpha$$

où $I_{C_1C_2C_3}^{C_i}$ désigne la composante C_i de l'image couleur I dans l'espace couleur $C_1C_2C_3$. La combinaison se fait avec deux composantes couleur provenant des espaces couleur parmi RVB , $L^*u^*v^*$ et HSL . α est le coefficient de pondération de la seconde image ($\alpha \in [0,1]$). L'image I_{CL} est l'image résultant de la combinaison linéaire de I et I' et représente leur fusion. Il nous reste à définir quels composantes couleur utiliser, comment évaluer la pertinence d'une fusion et comment choisir le coefficient α . Pour cela, nous allons utiliser les informations a priori pour apprendre comment fusionner les composantes couleur et construire une image monochromatique donnant une représentation optimale de la couleur. On notera $LC(I, I')$ la fusion de deux composantes couleur I et I' .

2.2.3.3.2. Apprentissage par sélection séquentielle de composantes couleur

Pour choisir et évaluer la fusion de deux composantes on utilise les informations a priori. La fusion de deux images nous donne l'image I_{LC} . Cette image est segmentée en utilisant la méthode de Fisher. Cette méthode classe les pixels d'une image en k classes. Les zones segmentées ont la propriété de minimiser la somme des variances intra classes. Une fois que l'image de fusion a été segmentée, les objets extraits sont comparés à la segmentation manuelle et un taux de classification β est calculé (en terme de pixels bien classifiés). β évalue l'image de fusion I_{LC} et donc la pertinence de l'image par rapport aux informations a priori, c'est-à-dire sa pertinence vis à vis de la segmentation. Ce critère d'évaluation sera noté par g . Nous venons de définir comment évaluer une image obtenue par fusion par rapport aux informations a priori. A présent, nous devons définir comment explorer l'espace de recherche. Pour n composantes couleur, il y a $1+2+\dots+(n-1)$ combinaisons possibles. L'espace de recherche étant trop grand pour être exploré complètement, nous proposons d'utiliser une sélection séquentielle des composantes couleur afin d'évaluer leurs degrés de pertinence. Nous commençons par une image vide I' et nous essayons toutes les combinaisons possibles de I' avec les autres composantes couleur parmi $\{R, V, B, L^*, u^*, v^*, H, S, L\}$. Pour chaque

combinaison linéaire, toutes les valeurs du coefficient α sont testées (avec un pas de 0,1). Chaque combinaison est évaluée selon le critère que nous avons précédemment défini (la valeur de β). La meilleure combinaison de composantes (notée I_{Best}) peut être retenue en cherchant à maximiser β et l'image qui a servi est éliminée des composantes couleurs utilisables. L'algorithme itère ensuite : I est remplacé par I_{Best} et la sélection continue jusqu'à ce qu'aucune amélioration ne soit obtenue dans la valeur de β . Cet algorithme est réalisé sur l'ensemble des images représentatives qui ont été utilisées pour extraire les informations a priori. Cela implique d'évaluer chaque combinaison de composantes couleur pour chaque image de la base d'images représentatives. Une combinaison linéaire est évaluée pour chacune des images. L'évaluation globale est définie comme la moyenne des évaluations sur les images de la base. Cet algorithme est utilisé pour trouver une combinaison optimale de composantes couleur, cette composante donnant une nouvelle représentation de la couleur relativement aux informations a priori. Une fois que la formule de combinaison linéaire a été définie, on peut l'utiliser pour obtenir une représentation des images couleur afin de les segmenter.

2.2.3.3.3. Résultats

Nous avons appliqué notre algorithme à des images de microscopie. Quatre images représentatives du problème de segmentation ont été segmentées manuellement. Afin de comparer l'effet produit par les images retenues pour constituer la base sur la représentation obtenue, nous avons appliqué l'algorithme sur chacune des images indépendamment et sur toute la base. On peut constater que les composantes couleur V, v^* et S sont toujours présentes dans les formules (Tableau 4). Ainsi ces composantes couleur contiennent des informations qui sont directement liées aux informations a priori et elle peuvent être définies comme pertinentes. De plus, la recherche d'un espace de représentation doit se faire sur une base d'images représentatives car les formules de combinaison sont différentes suivant les images utilisées.

Numéro de l'image	Combinaison
1	$((0,7*V)*0,8+0,2*S)*0,8+0,2*v^*$
2	$((((0,1*L^*)*0,7+0,3*v^*)*0,8+0,2*V)*0,9+0,1*S)$
3	$((0,1*V)*0,8+0,2*S)*0,1+0,9*v^*$
4	$((((0,1*V)*0,6+0,4*v^*)*0,8+0,2*S)*0,9+0,1*L^*)$
Tous	$((0,42*V)*0,83+0,17*S)*0,63+0,37*v^*$

Tableau 4. Les formules de combinaisons selon les images utilisées dans la base.

Afin de comparer notre stratégie de fusion de composantes couleur, nous avons comparé le résultat de la segmentation par la méthode de Fisher sur notre image obtenue par combinaison avec ceux obtenus sur la composante bleue et sur la première composante de la transformée de Karhunen-Loève (qui est l'axe le plus discriminant) que nous appellerons X_1 . Les régions extraites par la segmentation sont plus précises en utilisant notre image de combinaison (Figure 22). Nous pouvons donc dire que la représentation couleur que nous avons obtenue est adaptée pour segmenter de nouvelles images. L'utilisation de la fusion et d'informations a priori permet d'obtenir une représentation de la couleur qui s'avère utile pour segmenter des images.

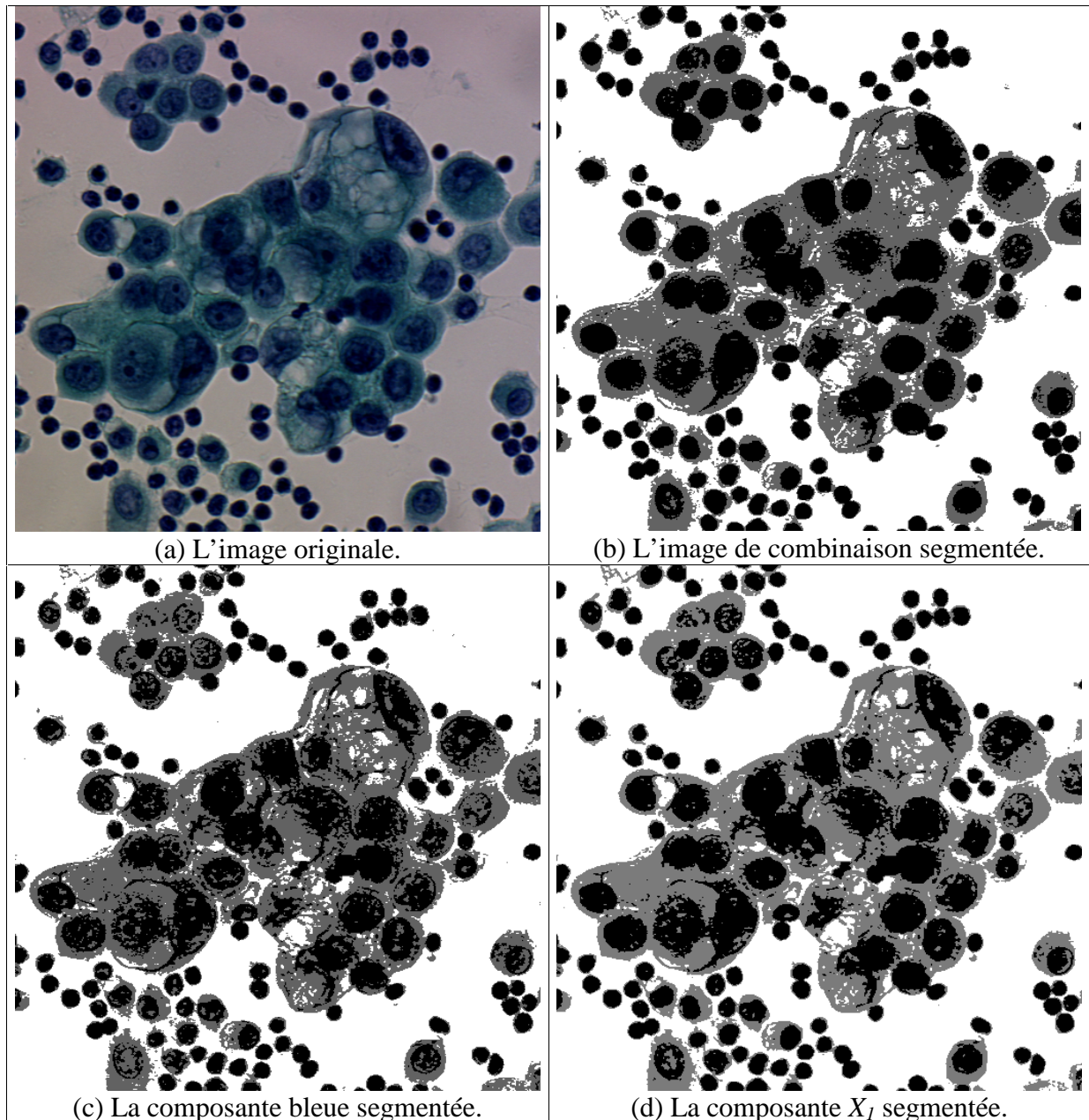


Figure 22. Les régions obtenues avec l'image de combinaison, la composante bleue et la composante X_7 .

2.2.3.4. Bilan

Au vu des études des paragraphes précédents, plusieurs remarques s'imposent :

- Les espaces couleur uniformes permettent des mesures plus fines que les autres espaces mais étant donné que certains espaces moins performants donnent des résultats satisfaisants, il n'y a pas lieu de les utiliser.

- Par contre pour certains cas d'étude où la sensibilité d'étude est particulièrement fine (zones homogènes, zones de dégradé continu etc.), il est nécessaire d'utiliser des espaces uniformes car leur dispersion spatio-colorimétrique est meilleure.
- Le choix d'un espace couleur en vue de la résolution d'un problème d'analyse d'image nécessite une connaissance a priori de l'information recherchée. Nous avons vu une méthode qui permet de déterminer quel espace couleur est le mieux adapté à un processus de segmentation. Cette méthode permettra d'utiliser l'information couleur pour un problème donné puisque l'espace couleur qui donne le meilleur contraste est celui qui caractérise le mieux la dispersion des couleurs vis-à-vis du problème traité. Cette méthode apparaît donc comme une méthode utile pour choisir un espace couleur dans un processus de segmentation d'images couleur. Nous précisons encore une fois que la méthode de choix d'un espace couleur est liée à notre méthode de segmentation couleur et à l'application étant donné la forte utilisation d'informations a priori.

Dans la suite nous n'utiliserons que la méthode de choix de l'espace couleur basée sur le contraste.

2.2.4. Gradient couleur

Le calcul du gradient dans un espace couleur peut se faire sur une seule, sur deux ou trois composantes. Cependant l'utilisation d'un gradient calculé sur les trois composantes à la fois tiendra mieux compte de la dispersion des couleurs et sera mieux approprié à un processus de type croissance de régions. Une image couleur peut être vue comme une fonction

$F : R^2 \rightarrow R^m$ où m est le nombre de composantes. Di Zeno [Di Zeno 83] a montré comment étendre le concept de gradient aux images multi-composantes. Ce gradient est applicable à des images contenant un nombre arbitraire de composantes sous deux conditions, la première étant que les informations dans les différentes bandes doivent être corrélées et la deuxième que le niveau de bruit doit y être équivalent. L'idée principale consiste à chercher une mesure du contraste dans l'image. La première dérivée directionnelle de f dans la direction $\vec{u} = (u_1, u_2)$ est elle-même un vecteur à m composantes. On cherche la direction \vec{n} qui correspond à la plus grande pente de la fonction, c'est-à-dire telle que $\left\| \frac{\partial f}{\partial \vec{n}} \right\|$ soit maximal.

Dans le cas de la norme euclidienne nous avons :

$$D_{\vec{u}} = \left\| \frac{\partial f}{\partial \vec{n}} \right\| = E u_1^2 + 2F u_1 u_2 + G u_2^2 \quad \text{où}$$

$$E = \sum_{i=1}^m \left(\frac{\partial f_i}{\partial x} \right)^2, \quad F = \sum_{i=1}^m \frac{\partial f_i}{\partial x} \frac{\partial f_i}{\partial y}, \quad G = \sum_{i=1}^m \left(\frac{\partial f_i}{\partial y} \right)^2$$

$D_{\vec{u}}$ définit donc une mesure du gradient au carré dans une direction \vec{u} quelconque. Le contraste maximal, λ_{\max} , correspond à la plus grande valeur propre de $J^t J$, avec J la matrice Jacobien de f tandis que la direction de cette variation maximale, \vec{n} correspond au vecteur propre associé :

$$\lambda_{\max} = \frac{E + G + \sqrt{(E - G)^2 + 4F^2}}{2} \quad \text{et} \quad \vec{n} = \pm(F, \lambda_{\max} - E).$$

Donc λ_{\max} et \vec{n} correspondent à la définition de module au carré et de la direction du gradient multispectral. Cependant le fait que les différentes composantes doivent être

corrélées est une importante limitation. En effet, un changement d'espace couleur permet, selon l'espace, de décorréler les composantes, si bien que le gradient de Di Zenzo n'est plus applicable. Il est donc préférable d'utiliser un autre gradient couleur que l'on puisse appliquer dans n'importe quel espace couleur.

L'opérateur gradient que nous avons utilisé est un gradient couleur, il est donc défini sur un espace couleur et utilise les trois composantes données d'un espace couleur. Nous avons utilisé le gradient suivant :

$$\|\nabla I_{C_1 C_2 C_3}(x, y)\| = \sqrt{\left(\|\nabla I_{C_1}(x)\| + \|\nabla I_{C_2}(x)\| + \|\nabla I_{C_3}(x)\|\right)^2 + \left(\|\nabla I_{C_1}(y)\| + \|\nabla I_{C_2}(y)\| + \|\nabla I_{C_3}(y)\|\right)^2}$$

Les filtres utilisés pour calculer le gradient suivant les directions x et y sont respectivement FX et FY définis pour un voisinage huit (i.e. on a une trame carrée) :

$$FX = \frac{1}{12} \begin{pmatrix} 1 & 4 & 1 \\ 0 & 0 & 0 \\ -1 & -4 & -1 \end{pmatrix} \quad FY = \frac{1}{12} \begin{pmatrix} 1 & 0 & -1 \\ 4 & 0 & -4 \\ 1 & 0 & -1 \end{pmatrix}$$

Pour avoir une image de gradient indépendante de l'orientation, il faut faire des convolutions avec tous les filtres déduits les uns des autres par rotation. Ces convolutions ne sont pas effectuées en série mais en parallèle, c'est à dire que l'étape intermédiaire consistera à former, à partir de l'image initiale, des images gradients $\nabla_{\theta} I_{C_1 C_2 C_3}(p)$ pour chaque direction (8 pour notre trame carrée). A partir de là, l'image gradient sera obtenue en prenant comme valeur le maximum des valeurs de $\nabla_{\theta} I_{C_1 C_2 C_3}(p) : \nabla I_{C_1 C_2 C_3}(p) = \sup_{\theta} |\nabla_{\theta} I_{C_1 C_2 C_3}(p)|$. Plusieurs illustrations d'images de gradient sont données par la suite.

2.3. Expérimentations de notre approche de segmentation couleur

Nous disposons à présent d'une méthode générale de segmentation d'images couleur. Notre stratégie d'optimisation définit comment traiter les choix importants relatifs à l'application de la ligne de partage des eaux couleur pour un problème particulier. Nous allons voir sur deux applications, des stratégies complètes d'extraction d'objets qui utilisent la ligne de partage des eaux couleur. Pour chaque application nous exposerons la stratégie d'analyse d'image puis nous détaillerons chacun de ses points. Pour chacune des applications la résolution est guidée par l'information couleur. C'est cette information qui caractérise principalement nos images et c'est cette information qui guidera les différentes stratégies de segmentation (à travers la ligne de partage des eaux ainsi que d'autres opérations de traitement d'images).

2.3.1. Notations

Avant de développer en détail chacune des stratégies de segmentation pour les différentes applications, précisons les notations que nous allons utiliser.

- ✗ I désigne une image en niveaux de gris,
- ✗ $I_{C_1 C_2 C_3}$ désigne une image couleur dans l'espace couleur $C_1 C_2 C_3$,

- ✗ $I_{C_1C_2C_3}^{C_1}$ désigne la composante C_1 de l'image couleur $I_{C_1C_2C_3}$,
- ✗ $S(I)$ désigne le seuillage d'une image I ,
- ✗ $E^B(I)$ désigne l'érosion morphologique d'une image I avec l'élément structurant B ,
- ✗ $D^B(I)$ désigne la dilatation morphologique d'une image I avec l'élément structurant B ,
- ✗ $O^B(I)$ désigne l'ouverture morphologique d'une image I avec l'élément structurant B ,
- ✗ $I_1 \cap I_2$ désigne l'intersection morphologique entre deux images binaires I_1 et I_2 ,
- ✗ $Exp(I)$ désigne le lissage exponentiel d'une image I ,
- ✗ $\nabla(I_{C_1C_2C_3})$ désigne le gradient couleur de l'image couleur $I_{C_1C_2C_3}$,
- ✗ $Cdist(I)$ désigne la distance de chanfrein d'une image binaire I par un filtre 5x5,
- ✗ $UES(I)$ désigne l'ensemble des érodés ultimes d'une image I ,
- ✗ $LAB(I)$ désigne l'étiquetage d'une image binaire I ,
- ✗ $WTS(I_P, I_M)$ désigne la ligne de partage des eaux d'une image I_p avec une image de marqueurs I_M ,
- ✗ $CWTS(I_P, I_M, I_{C_1C_2C_3})$ désigne la ligne de partage des eaux couleur d'une image couleur $I_{C_1C_2C_3}$ avec une image de gradient I_p , une image de marqueurs I_M et le paramètre α .

2.3.2. Cytologie des séreuses

2.3.2.1. Problématique et description des images

Les images que nous devons segmenter pour cette application sont des images de la cytologie des séreuses. Sur ces images on trouve des cellules qui doivent être isolées. Pour cela il faut extraire le cytoplasme et le noyau de chacune des cellules présentes. Le cytoplasme permet d'avoir une information de contexte (car les cellules peuvent être aussi bien isolées qu'en amas) et le noyau permet d'apprécier la malignité d'une cellule. Les deux informations permettent ensuite par un processus de classification de différencier les différents types cellulaires. Avant de segmenter nos images il faut décrire précisément les types de fond et les types d'objets que nous pouvons rencontrer. On distingue trois types de fond : les fonds clairs homogènes, les fonds hémorragiques (présence de globules rouges en plus des cellules) et les fonds muco-protéiniques (présence de mucus et de protéines colorés en vert). Sur une image on trouve des cellules isolées, accolées, se chevauchant, en amas ou en papilles. Une cellule est soit un globule rouge et est coloré en rouge, soit une cellule dont le noyau est coloré en bleu et le cytoplasme en vert. La stratégie de segmentation que nous avons mis en place doit extraire les cellules présentes quelles que soient leurs configurations spatiales (de cellule isolée à papille : Figure 23) à l'exception des globules rouges qui ne présentent aucun intérêt d'interprétation.

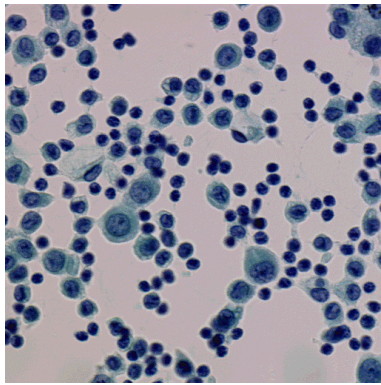


Figure 23.a. Un fond homogène.

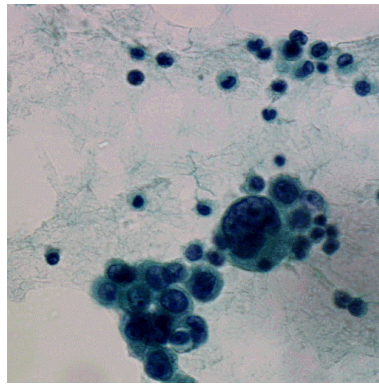


Figure 23.b. Un fond muco-protéinique

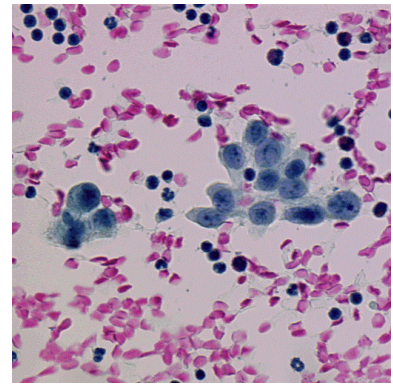


Figure 23.c. Un fond hémorragique.

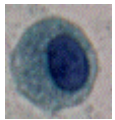


Figure 23.d

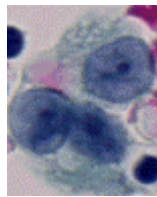


Figure 23.e

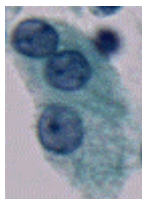


Figure 23.f

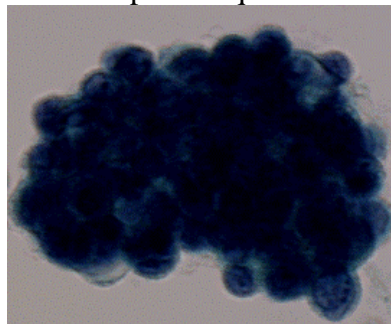


Figure 23.g. Une papille avec chevauchement des noyaux.

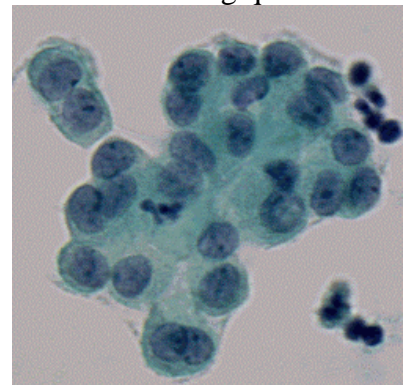


Figure 23.h. Une papille avec noyaux accolés.

Figure 23.d.e.f. Cellule isolée, cellules se chevauchant, cellules accolées.

Figure 23. Les différentes configurations spatiales des cellules.

2.3.2.2. Stratégie de segmentation

Notre stratégie de segmentation doit isoler le noyau et le cytoplasme de toutes les cellules présentes sur une image à l'exception des globules rouges. Nous procédons en six étapes successives. La première étape consiste à déterminer si il y a présence ou non de noyaux sur l'image à analyser. Si des cellules sont présentes, on procède à l'extraction des globules rouges afin de les éliminer. Les étapes suivantes concernent l'extraction des cytoplasmes et des noyaux et sont réalisées par notre méthode de ligne de partage des eaux couleur. Une fois les cellules extraites, les noyaux accolés sont séparés et finalement les cytoplasmes sont divisés afin de ne contenir qu'un seul noyau ou amas de noyaux. Dans une ultime étape les noyaux touchant le bord de l'image sont éliminés. Cette stratégie est ascendante, nous déterminons des informations globales sur les images à travers les premières étapes, puis nous portons notre attention de façon de plus en plus précise sur les objets à extraire. Les objets définitifs sont des cellules définies par un cytoplasme et un noyau. Nous verrons par la suite que ce noyau pourra en fait être un noyau seul ou un amas de noyaux. La stratégie est exprimée sur la Figure 24 par un graphe suivant la définition donnée dans [Clouard 99]. Chaque boîte représente un opérateur (une opération de traitement d'images) ou une structure de contrôle (un if par exemple). Les flèches noires représentent le flux des

données (les images), les flèches bleues représentent le flux de contrôle (l'enchaînement des opérateurs) et les flèches rouges représentent le flux des valeurs numériques et les flèches vertes le flux des images de référence.

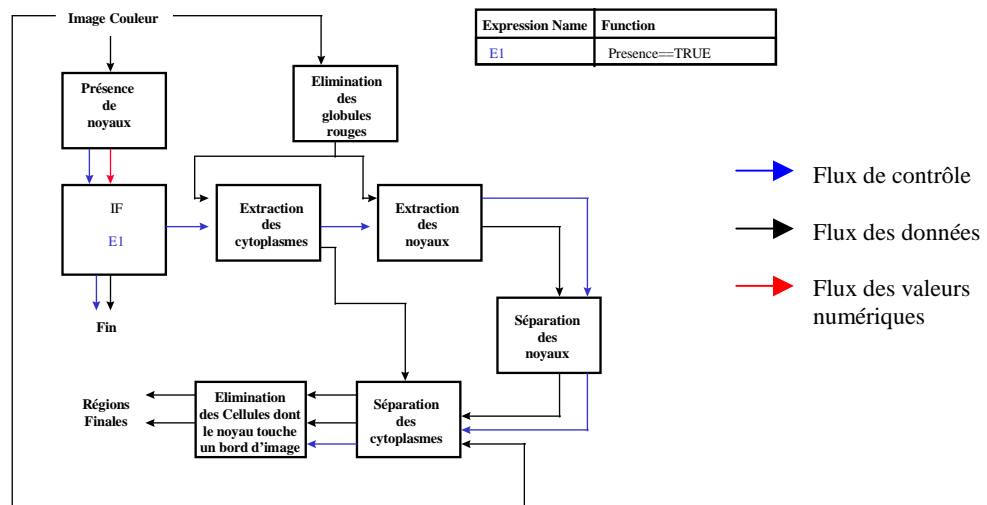


Figure 24. La stratégie de segmentation d'images de cytologie des séreuses.

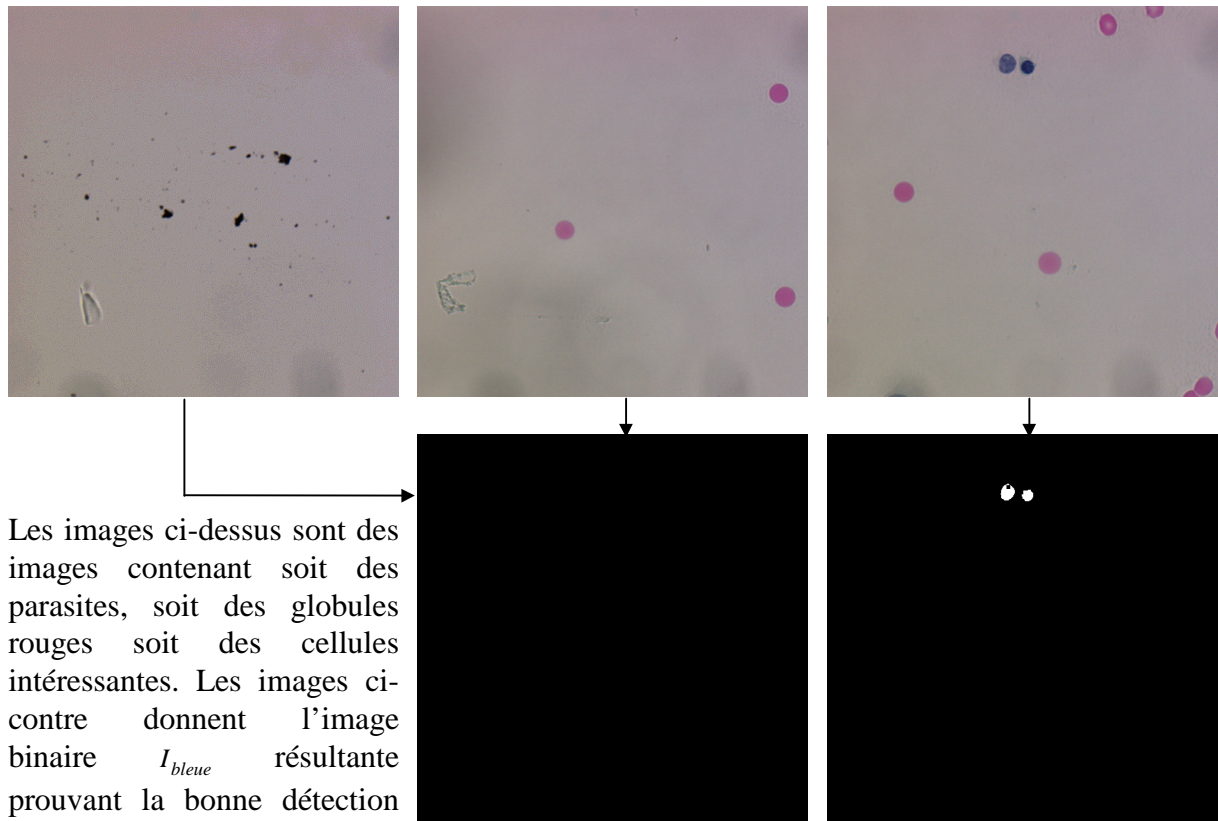
2.3.2.3. Détection des cellules intéressantes

Avant de segmenter une image il faut déterminer si des cellules intéressantes sont présentes sur celle-ci. Les cellules présentant un intérêt sont les cellules possédant un noyau bleu. Le problème est donc réduit à la détection des objets bleus présents sur une image. Cette étape servant principalement à réduire le temps de calcul en ne traitant que les images présentant un intérêt particulier, nous avons mis en place une solution simple, rapide et très efficace.

La détection d'objets bleus peut s'avérer assez difficile si l'on travaille dans un espace couleur non approprié. Or ici l'information a priori que représente la détection d'objets bleus peut se traduire directement à travers l'expression de la teinte. Un simple seuillage de la teinte dans l'espace HSL permet de déterminer si l'image contient du bleu ou non. En prenant deux seuils S_1 et S_2 définissant les bornes supérieures et inférieures d'un intervalle de teinte, on peut définir un seuillage fixe de façon à ce que l'intervalle $[S_1, S_2]$ ne contienne que des couleurs dont la teinte s'approche du bleu. Nous avons choisi les seuils S_1 et S_2 de façon à ce que toutes les couleurs de teinte bleue soient contenues dans cet intervalle (même si l'on déborde sur les couleurs de teinte rouge ou verte). Comme cette sélection n'est pas très précise deux érosions successives sont calculées sur l'image binaire résultant du seuillage. L'image résultante I_{bleue} est exprimée par :

$$I_{bleue} = E^{\square} \left(E^{\square} \left(S \left(I_{HSL}^H \right) \right) \right) \text{ avec } S \left(I_{HSL}^H(p) \right) = \begin{cases} 255 & \text{si } I_{HSL}^H(p) \in \left[\pi + \frac{\pi}{8}, \frac{3\pi}{2} \right] \\ 0 & \text{sinon} \end{cases}$$

Si L'image binaire finale I_{bleue} est égale à l'ensemble vide i.e. si $\nexists p(x, y) \in I_{bleue}$ tel que $I_{bleue}(p) > 0$ alors aucun noyau bleu n'est présent sur l'image et l'image ne sera pas traitée (Figure 25).



Les images ci-dessus sont des images contenant soit des parasites, soit des globules rouges soit des cellules intéressantes. Les images ci-contre donnent l'image binaire I_{bleue} résultante prouvant la bonne détection des cellules intéressantes.

Figure 25. Exemples d'extraction des cellules intéressantes.

2.3.2.4. Elimination des globules rouges

Nous avons pu déterminer la présence de cellules intéressantes caractérisées par leur noyau bleu. Cependant sur certaines images présentant un fond hémorragique, certaines cellules (les globules rouges) peuvent gêner des traitements ultérieurs. Parfois, seules quelques cellules intéressantes sont présentes et par contre beaucoup de globules rouges apparaissent sur l'image (Figure 23.c). Segmenter ces cellules de la même façon que les cellules intéressantes et procéder à leur reconnaissance par une méthode de classification ne présente donc aucun intérêt. Nous nous sommes donc attachés à définir une méthode procédant à l'extraction des globules rouges sur une image en vue de leur élimination. Comment procéder ? Si l'on considère les informations couleur caractérisant un globule rouge, la principale caractéristique (et la seule) permettant de les identifier est leur couleur rouge. Une méthode simple procédant à l'extraction des objets rouges d'une image devrait donc suffire. Fort de la méthode mise en place au paragraphe précédent nous avons mis en œuvre une méthode identique. Le résultat de ces opérations est un masque I_{rouge} divisant l'image en deux régions : les globules rouges et le reste de l'image (les cellules et le fond). Par ce masque nous réduisons considérablement le nombre de cellules à segmenter sans pénaliser la future reconnaissance des cellules intéressantes car l'intervalle de variation est défini de façon à ne pas éliminer les parties de globules rouges chevauchant une cellule. I_{rouge} s'exprime par :

$$I_{rouge} = O^{\square}(s(I_{HSL}^H)) \text{ avec } s(I_{HSL}^H(p)) = \begin{cases} 0 & \text{si } I_{HSL}^H(p) \in \left[0, \frac{\pi}{3}\right] \cup \left[\frac{3\pi}{2}, 2\pi\right] \\ 255\sin\theta & \text{sinon} \end{cases}$$

Ce masque sera utilisé lors des étapes suivantes pour déterminer les marqueurs dans le processus de la ligne de partage des eaux couleur. Un exemple de masque est donné sur la Figure 26. On remarquera que la masque a été conçu pour éliminer également certains débris cellulaires que l'on retrouve souvent sur les images : les squames cornées anuclées (Figure 27).

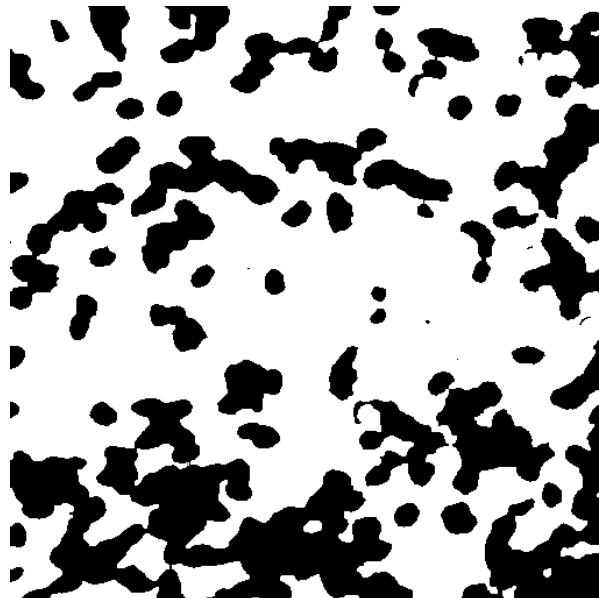
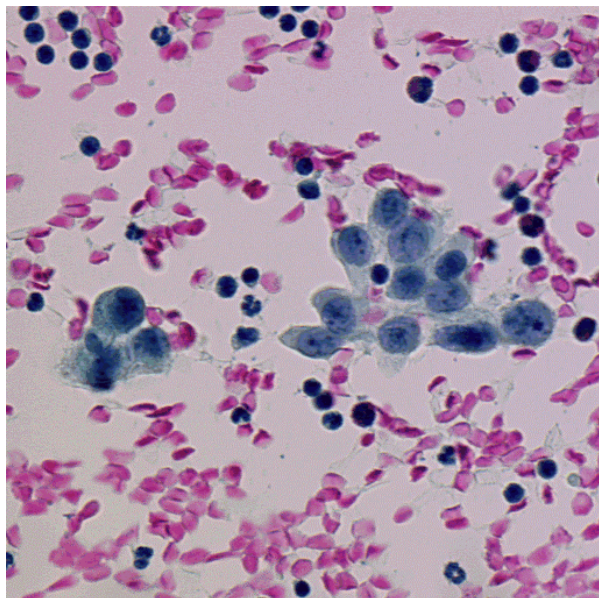


Figure 26.a. Une image à fond hémorragique. Figure 26.b. Le masque I_{rouge} correspondant.

Figure 26. Une image de cellules contenant des globules rouges et le masque correspondant.

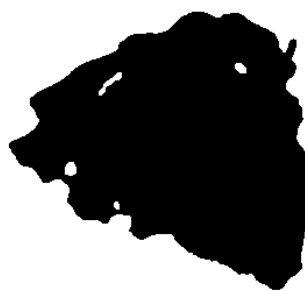
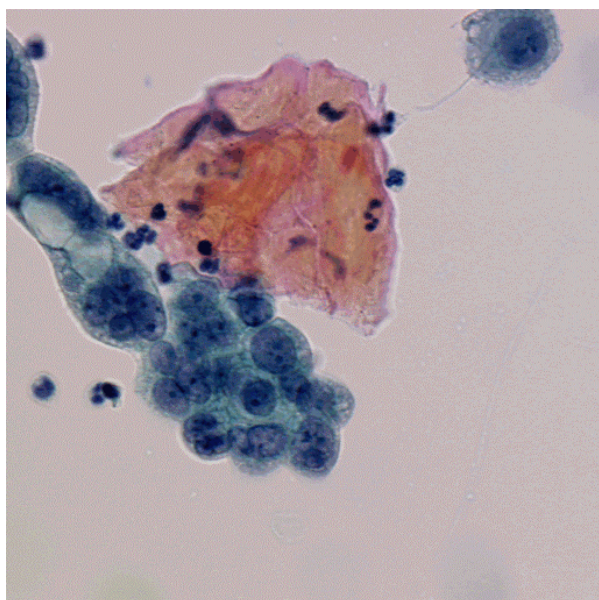


Figure 27.a. Une image contenant un débris cellulaire. Figure 27.b. Le masque I_{rouge} correspondant.

Figure 27. Une image de cellules en amas accolées avec un débris cellulaire et le masque correspondant.

2.3.2.5. Extraction des régions cytoplasmiques

Pour extraire les régions cytoplasmiques des images, nous avons utilisé la ligne de partage des eaux couleur. Nous allons voir chacune des étapes de la stratégie d'optimisation de la fonction d'agrégation pour les images traitées.

1. Le choix de l'espace couleur a été réalisé sur des images représentatives des différentes cellules et des différents fonds pouvant être rencontrés. On trouve également sur ces images les diverses configurations spatiales possibles des cellules ou amas de cellules. Sur ces images nous avons appliqué notre méthode servant à déterminer quel espace est le mieux adapté à un processus de segmentation. Les résultats sont donnés dans le Tableau 5. L'espace couleur présentant le meilleur contraste pour toutes les images est l'espace *RVB*. Cet espace sera donc utilisé pour la L.P.E couleur.
2. La métrique est une métrique euclidienne.
3. Le calcul du gradient est effectué dans l'espace *RVB* puisque c'est cet espace qui a été choisi en 1 (Figure 28.b).
4. Pour déterminer les marqueurs, les informations a priori que nous pouvons utiliser sont la couleur verte du cytoplasme ainsi que la transition franche entre fond et cytoplasme quel que soit le type fond. Cette dernière propriété se traduit directement par l'information de luminance. Pour que l'extraction des marqueurs se fasse de façon précise et reproductible, l'histogramme de l'image de luminance doit avoir sensiblement la même forme quelles que soient les images à segmenter. Nous avons choisi la luminance de l'espace couleur *HSL* car elle satisfaisait à ces exigences. Après un lissage par un filtre exponentiel, l'image présente des transitions plus franches et moins bruitées entre les cytoplasmes et le fond. Le seuil servant au seuillage de l'image de luminance lissée est fixé comme étant le maximum de la variance interclasse et de la vallée de l'histogramme. Ce seuil permet d'extraire rapidement et de façon sûre les cytoplasmes des cellules qu'elles soient isolées ou en amas. Pour éliminer les globules rouges, il suffit de calculer un intersection morphologique entre les marqueurs et le masque des globules rouges. L'image résultante est une image binaire contenant seulement les marqueurs cytoplasmiques. Les marqueurs sont érodés plusieurs fois afin d'augmenter leur zone d'influence pour obtenir une meilleure précision dans la définition des régions cytoplasmiques. L'image des marqueurs cytoplasmiques (Figure 28.a et c) s'exprime par : $I_{\text{CytoplasmeMarqueurs}} = I_{CM} = E^{\square} \left(I_{\text{rouge}} \cap S \left(\text{Exp} \left(I_{HSL}^H \right) \right) \right)$
5. Comme nous l'avons vu, la limite entre le fond et les cytoplasmes est bien marquée et l'information gradient est suffisante pour extraire les cytoplasmes des cellules. Nous avons donc fixé α à 1 afin de n'utiliser que l'information liée au gradient couleur dans le processus de croissance. L'image résultante obtenue (Figure 28.d) par la ligne de partage des eaux couleur est $I_{\text{Cytoplasmes}} = I_C = \text{CWTS}(\nabla(I_{RVB}), \text{LAB}(I_{CM}), I_{RVB})$.

Par cette méthode les cytoplasmes des cellules sont correctement segmentés (Figure 28) quelles que soient les cellules à segmenter mais cela ne sépare pas les amas de cytoplasmes qui présentent peu ou pas de différences d'intensité entre tous les cytoplasmes de l'amas.

Fond	<i>RVB</i>	<i>XYZ</i>	<i>L* u* v*</i>	<i>L* a* b*</i>	<i>HSL</i>	Chassery
Homogène	67,20	64,25	45,19	27,83	64,49	24,53
Hémorragique	87,59	84,45	60,61	34,53	85,23	31,33
Muco-protéinique	60,59	58,05	53,35	25,95	59,39	38,38
Moyenne	71,79	68,91	53,05	29,43	69,70	31,41

Tableau 5. Les contrastes suivant les différents fonds des images.

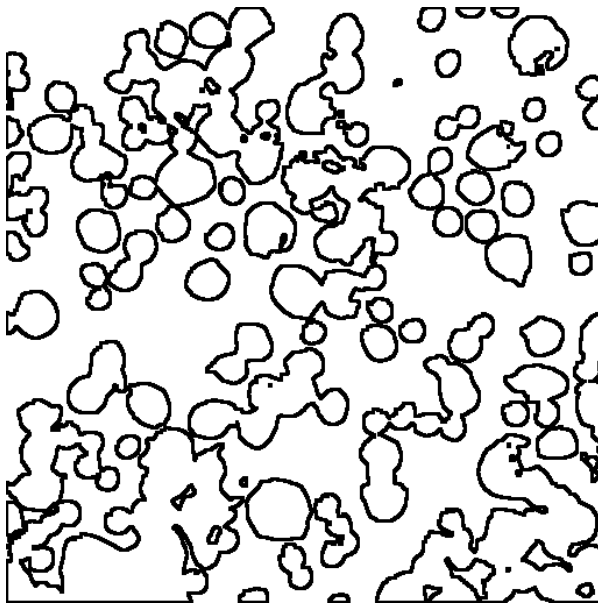
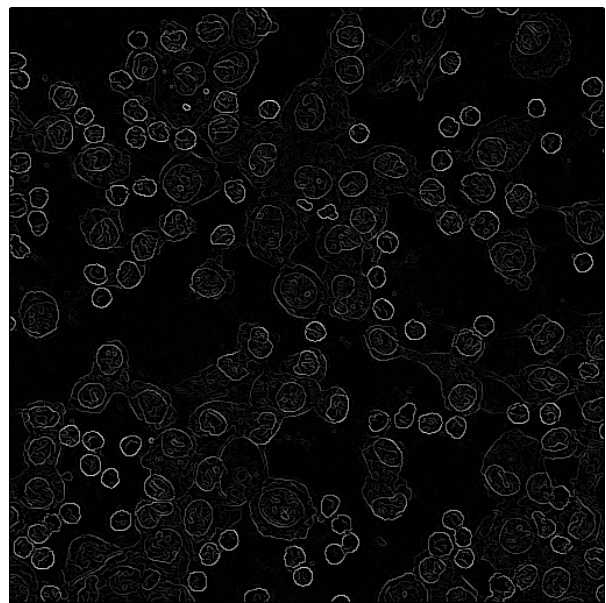
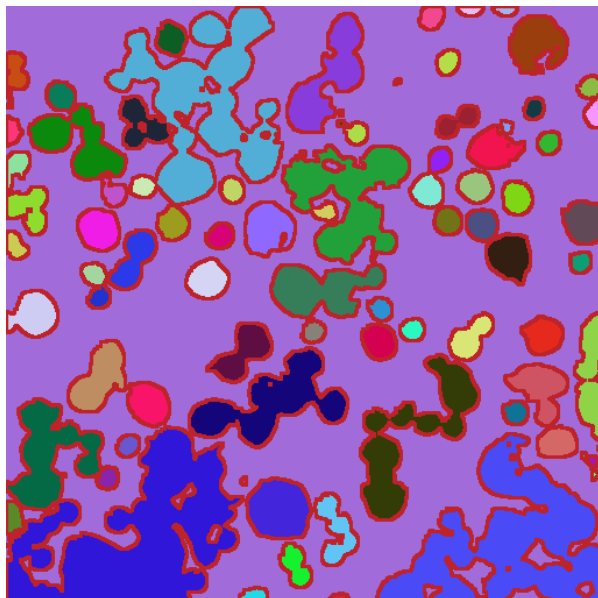
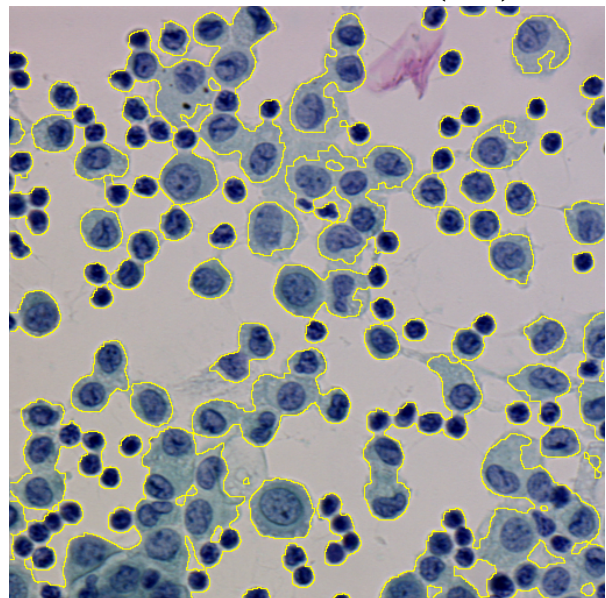
Figure 28.a. Les marqueurs (I_{CM}).Figure 28.b. Le gradient ($\nabla(I_{RVB})$).Figure 28.c. Les marqueurs étiquetés
($LAB(I_{CM})$).Figure 28.d. Les régions cytoplasmiques (I_C).

Figure 28. Les marqueurs (a), le gradient (b), les marqueurs étiquetés (c) et les régions cytoplasmiques en jaune (d).

La stratégie complète d'extraction des cytoplasmes est exprimée sur la Figure 29 par un graphe d'opérateurs.

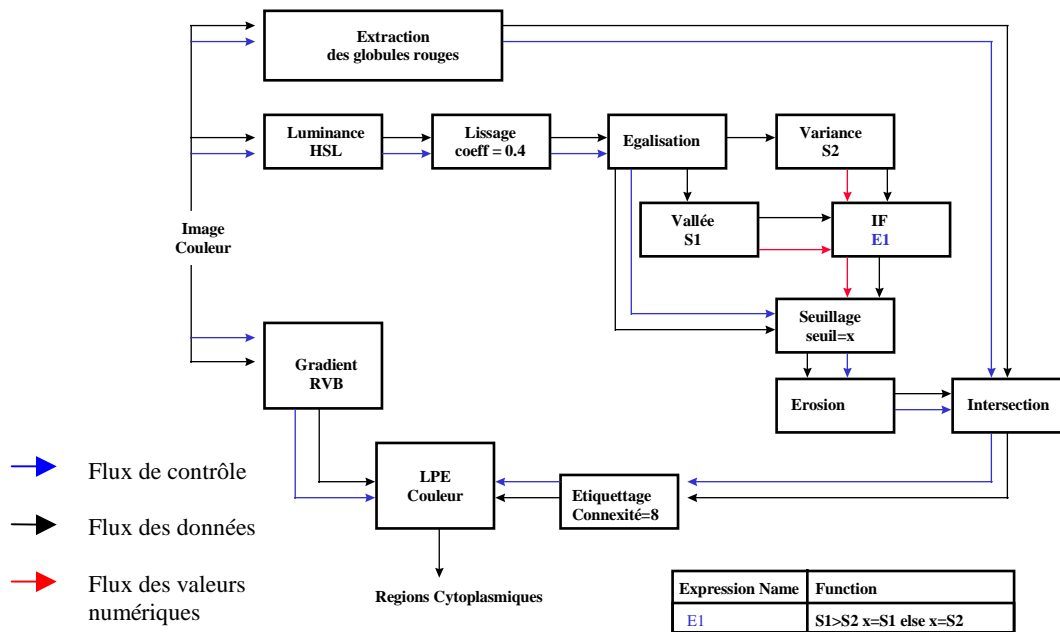


Figure 29. Graphe de la segmentation des cytoplasmes.

2.3.2.6. Extraction des régions nucléaires

Cette extraction des régions nucléaires est également basée sur la ligne de partage des eaux couleur. Voyons en détail la stratégie d'optimisation de la fonction d'agrégation.

1. Le choix de l'espace couleur se fait sur des images représentatives des noyaux des cellules pouvant être rencontrés. Comme nous l'avons vu précédemment les noyaux ont une extrême variabilité dans leur configuration spatiale et de leur couleur. Nos images contiennent suffisamment de cellules dont les noyaux sont représentatifs pour avoir une mesure du contraste valide et fiable. Le Tableau 6 donne les mesures de contraste pour chacun des espaces couleur. Nous avons pris trois types de cellules. Le type 1 correspond à une image contenant des cellules aux noyaux isolés ou accolés dont la couleur est bleu foncé, le type 2 à des cellules aux noyaux isolés ou en amas dont le noyau est très foncé et le type 3 à des cellules dont le noyau est isolé ou accolé dont le noyau est bleu très clair. L'espace *RVB* qui présente la meilleure moyenne du contraste est choisi.
2. La métrique est une métrique euclidienne.
3. Le calcul du gradient est effectué dans l'espace *RVB* puisque c'est cet espace qui a été sélectionné en 1 (Figure 31.b).
4. Pour trouver les marqueurs des noyaux il faut déterminer une composante connexe et une seule par noyau ou amas de noyaux. Quelles informations a priori doivent guider notre recherche de composantes couleur à combiner ? La difficulté posée par l'extraction des marqueurs réside dans l'extraction des noyaux clairs et des noyaux foncés ainsi que dans les faibles transitions couleur que l'on peut rencontrer entre les noyaux et les cytoplasmes. Il nous faut donc trouver deux images : l'une faisant

ressortir les noyaux clairs et l'autre les noyaux foncés, ces deux images accentuant également la transition du cytoplasme au noyau. Une combinaison de ces deux

images est donnée par : $I_S = \frac{(I_{RVB}^B - I_{RVB}^V) + I_{Luv}^{v^*}}{2}$. L'image I_S permet d'avoir une

meilleure distribution des niveaux de gris entre les noyaux et le reste de l'image : les noyaux sont contenus dans un intervalle de niveaux de gris plus restreint, ce qui permet une extraction plus sûre et reproductible de tous les noyaux qu'ils soient clairs ou foncés (Figure 30.c). La Figure 30.a donne l'exemple d'une image présentant des noyaux clairs et foncés et la Figure 30.b présente l'image I_S obtenue. Les noyaux clairs ont été accentués et la transition du cytoplasme au noyau est plus franche en comparaison d'une composante couleur seule (le bleu donné par la Figure 30.e et son histogramme par la Figure 30.d.)

L'extraction des marqueurs se fait par un seuillage automatique sur l'image I_S . Puisque l'histogramme de l'image I_S montre que les noyaux sont dans un intervalle de niveaux de gris restreint correspondant au pic principal de l'histogramme, l'obtention d'un seuil est plus sûre et reproductible. Le seuil est choisi à partir de deux seuils : la variance interclasse (S_1) et la vallée de l'histogramme (S_2 : le point d'inflexion de l'histogramme). Suivant les images à segmenter, il s'est avéré que le meilleur seuil était soit l'un soit l'autre suivant l'image à segmenter. Afin de choisir de façon quantitative et automatique le meilleur seuil, nous avons utilisé la propriété suivante : une image I_S seuillée est meilleure qu'une autre si les marqueurs contiennent moins de cytoplasme. Le cytoplasme étant caractérisé par sa couleur verte, le seuil choisi est celui qui minimise la quantité de vert incluse dans les marqueurs. Cette sélection s'exprime par :

$$\left\{ \begin{array}{l} T_1 = \left(\sum I_{RVB}^V(p) \right) / p \in I_S \\ \text{et } I_S(p) < S_1; \\ T_2 = \left(\sum I_{RVB}^V(p) \right) / p \in I_S \\ \text{et } I_S(p) < S_2; \\ \text{si } T_1 < T_2 \text{ alors } S = S_1 \text{ sinon } S = S_2 \end{array} \right.$$

Le seuil S obtenu a la propriété de permettre d'inclure tous les noyaux clairs et foncés. L'image I_S seuillée peut cependant contenir aussi des globules rouges (de façon similaire à l'obtention des marqueurs des cytoplasmes). On procède donc à une intersection avec le masque des globules rouges calculé au § 2.3.2.4. L'image finale représente les marqueurs nucléaires (Figure 31.a et c) et s'exprime par :

$$I_{\text{MarqueursNucléaires}} = I_{MN} = E^{\square} \left(I_{\text{mask}} \cap S(I_S) \right).$$

5. Les noyaux sont obtenus par la ligne de partage des eaux couleur et l'image finale s'exprime par : $I_{\text{Noyaux}} = I_N = CWTS \left(\nabla(I_{RVB}), LAB(I_{MN}), I_{RVB} \right)$. Le coefficient de pondération α est fixé à 0,5 ; ceci permet de prendre en compte à la fois les informations locales et globales. L'utilisation de ces deux informations donne des régions nucléaires plus précises car l'une d'entre elles ne suffirait pas. En effet d'une part la transition entre le cytoplasme et le noyau n'est pas assez franche pour avoir une grande amplitude du gradient et d'autre part quelques grandes magnitudes du gradient peuvent apparaître ailleurs qu'à la transition noyau cytoplasme (à cause des variations de texture à l'intérieur des noyaux). Ainsi l'utilisation du critère

d'homogénéité couleur permet d'obtenir des régions nucléaires uniformes et dont les contours sont bien déterminés (Figure 31.d).

On constate que les régions nucléaires obtenues par la ligne de partage des eaux sont correctes quelles que soient les configurations des cellules. Les noyaux obtenus sont soit isolés, soit en amas, soit en amas avec recouvrement.

Noyaux	<i>RVB</i>	<i>XYZ</i>	<i>L* u* v*</i>	<i>L* a* b*</i>	<i>HSL</i>	Chassery
Type 1	89,63	88,24	62,54	34,74	88,47	32,29
Type 2	68,57	67,81	67,31	47,05	68,61	64,42
Type 3	98,44	95,25	71,93	36,56	95,76	30,04
Moyenne	85,54	83,76	67,26	39,45	84,28	42,25

Tableau 6. Contrastes pour les différentes configurations de noyaux dans chaque espace couleur.

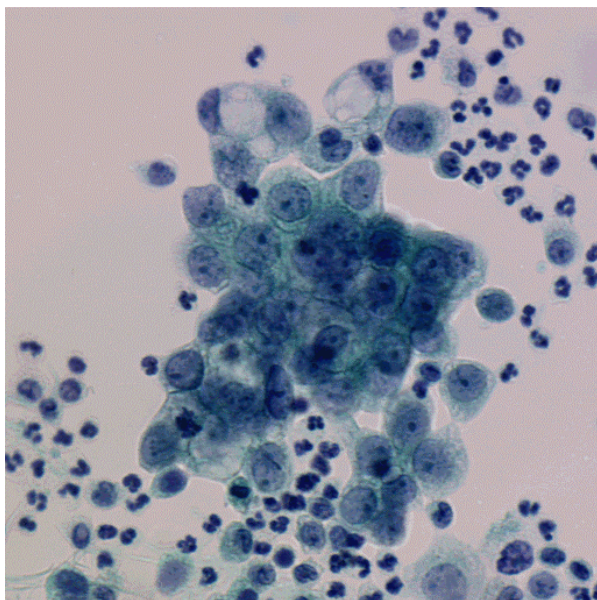


Figure 30.a. Une image couleur de cellules.

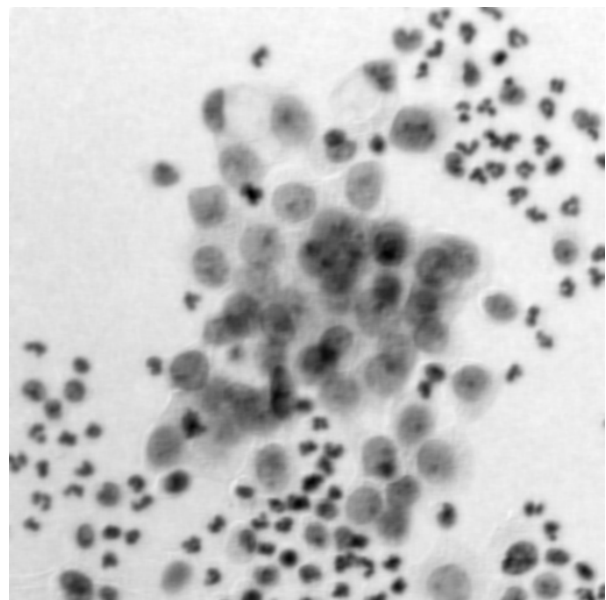


Figure 30.b. L'image I_S .

L'image Synthétique I_S permet d'accroître tous les noyaux (clairs et foncés) ainsi que la transition noyau-cytoplasme.

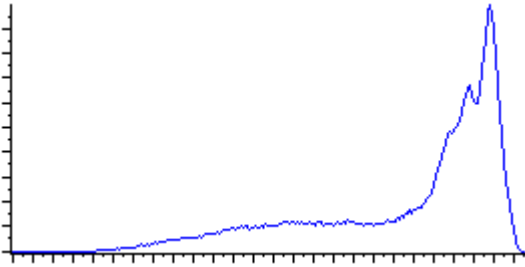


Figure 30.c. L'histogramme de l'image I_S .

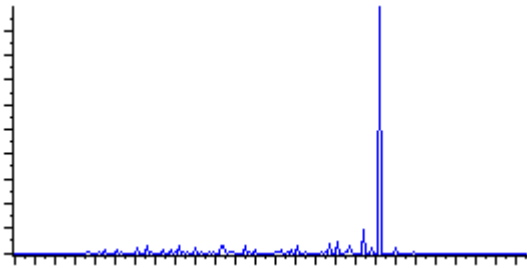


Figure 30.d. L'histogramme de la composante bleue.

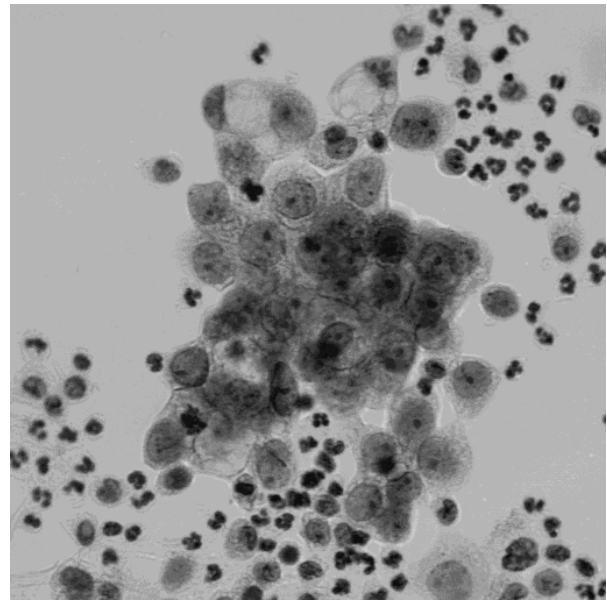


Figure 30.e. La composante bleue de la Figure 30.a.

Figure 30. Obtention de l'image I_S à partir d'une image couleur.

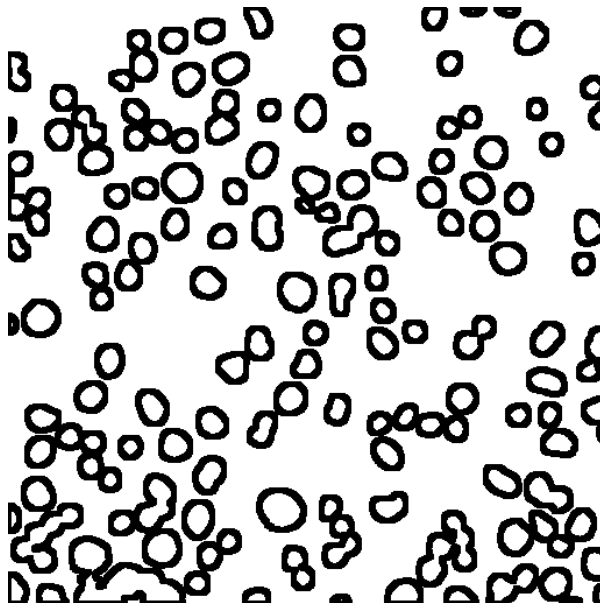


Figure 31.a. Les marqueurs (I_{MN}) .

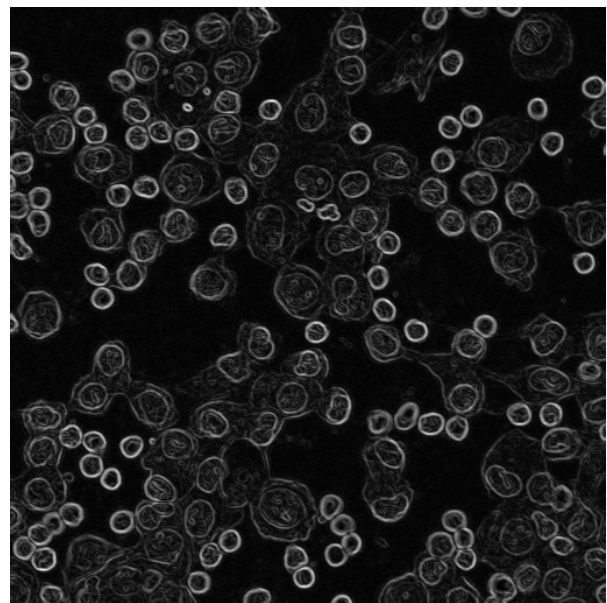


Figure 31.b. Le gradient $\nabla(I_{RVB})$.

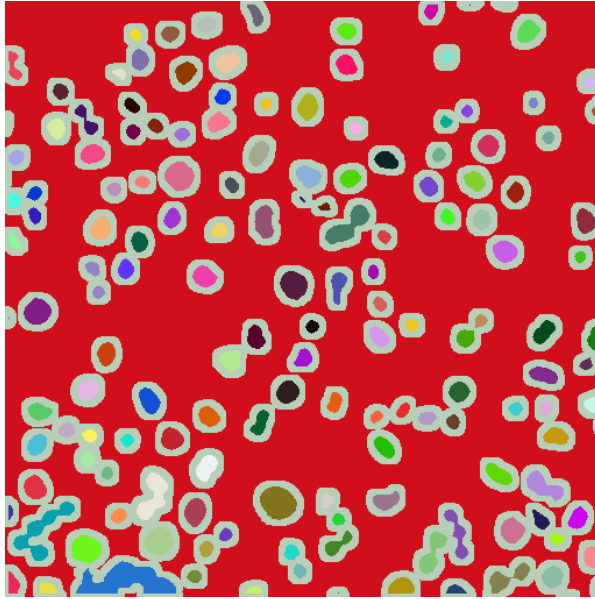


Figure 31.c. Les marqueurs étiquetés $LAB(I_{MN})$.

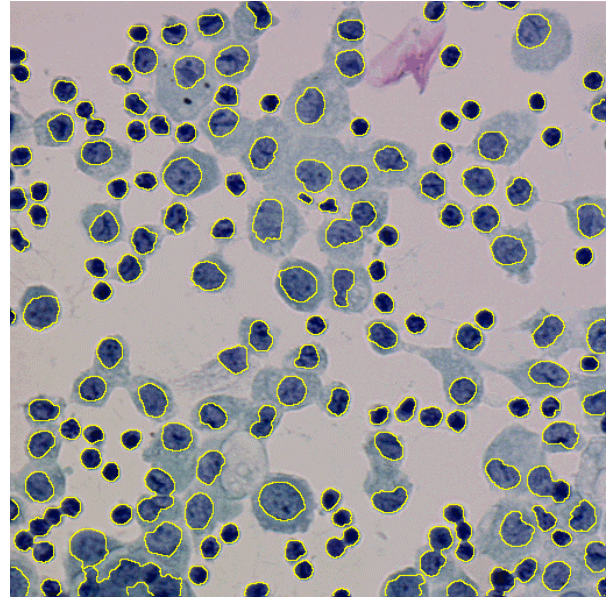


Figure 31.d. Les régions nucléaires (I_N).

Figure 31. Les marqueurs (a), le gradient (b), les marqueurs étiquetés (c) et les régions nucléaires en jaune (d).

La stratégie complète d'extraction des noyaux (Figure 31) est exprimée sur la Figure 32 par un graphe d'opérateurs.

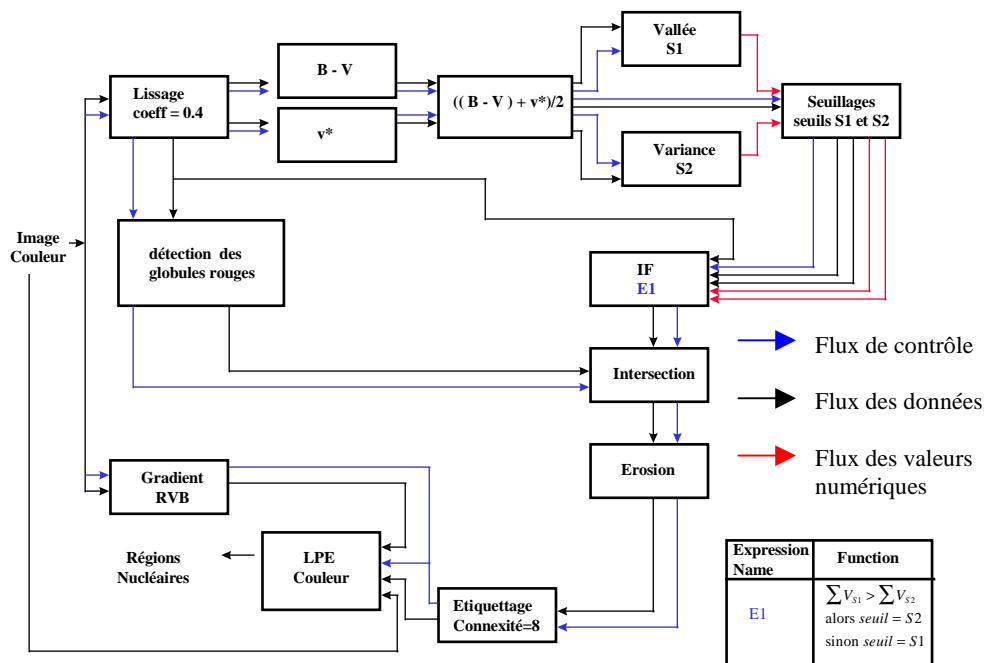


Figure 32. Graphe de la segmentation des noyaux.

2.3.2.7. Séparation des noyaux

Comme nous l'avons vu dans le chapitre précédent, les noyaux extraits ont été isolés dans une grande variété de configurations spatiales (isolés, accolés, en amas sans superposition, en amas avec superposition). Certaines de ces configurations atteignent un haut niveau de complexité dû aux chevauchement des noyaux. Mais dans certains cas des noyaux ont été segmentés comme n'étant qu'une seule région alors que ceux-ci peuvent être facilement divisés visuellement par un expert (Figure 33.a.) car le chevauchement est minime (Figure 33.b.). Evidemment pour les amas de noyaux présentant d'importants chevauchements des noyaux cette séparation n'apportera rien car l'interprétation directe des noyaux se chevauchant n'est pas réalisable par un expert. Nous avons donc choisi de mettre en place une méthode de séparation des noyaux qui séparera les noyaux simplement accolés. Ces noyaux « potentiellement » séparables sont reconnaissables à leur forme en particulier grâce à la présence de concavités dans leur bordure nucléaire. Nous allons donc utiliser cette propriété pour séparer les noyaux accolés. Une façon d'exprimer les variations dans la courbure des contours des noyaux est de calculer, pour chaque point d'un noyau, la distance à la bordure nucléaire correspondante. Les maxima de cette fonction distance se situent au centre de chaque noyaux et les minima sont les points reliant les concavités des noyaux accolés (Figure 33.d.). Une méthode simple pour séparer les noyaux accolés est d'utiliser une ligne de partage des eaux sur l'opposé de l'image de distance avec comme marqueurs l'ensemble des érodés ultimes de l'image de distance. La fonction distance utilisée est une distance de chanfrein (Figure 33.c.). L'ensemble des érodés ultimes est obtenu par une reconstruction morphologique de taille 2 sur l'image de distance (Figure 33 (d) et (e)). Les marqueurs fournissent donc une composante connexe pour chaque noyau. Par ce procédé de séparation des noyaux, les amas de noyaux sont séparés et les noyaux qui sont obtenus sont soit des noyaux isolés (Figure 33 (f)) soit des amas de noyaux se chevauchant (Figure 34 (a) et (b)). L'image finale des régions nucléaires s'exprime par :

$$I_{\text{NoyauxSéparés}} = I_{NS} = WTS\left(Cdist(I_N), LAB\left(UES\left(Cdist(I_N)\right)\right)\right).$$

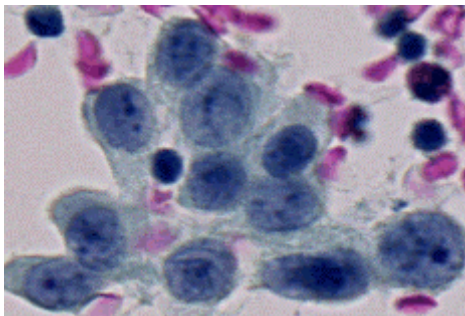


Figure 33.a. Noyaux accolés.

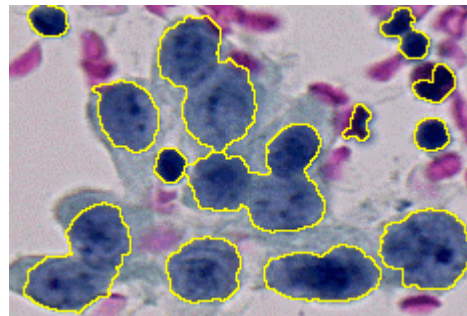


Figure 33.b. Noyaux accolés segmentés.



Figure 33.c. La fonction distance.



Figure 33.d. Les érodés ultimes.

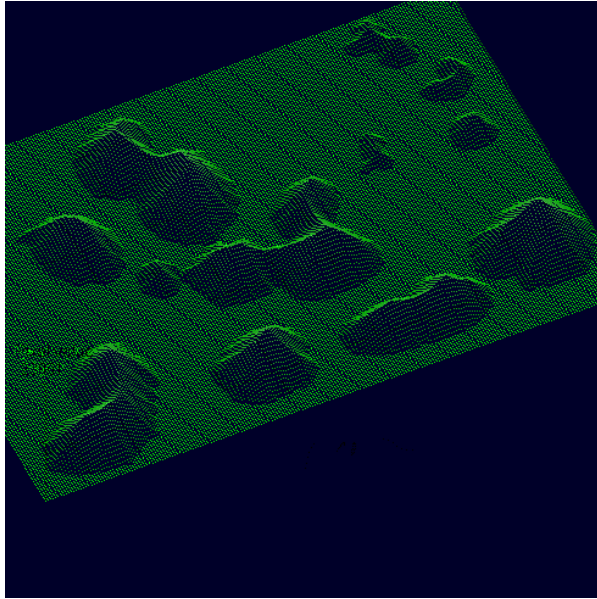


Figure 33.e. Vue 3D de la Figure 17.c.

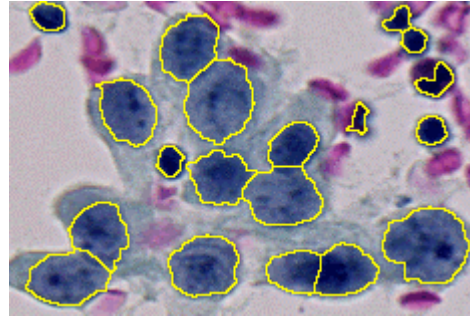


Figure 33.f. Les noyaux séparés.

Figure 33. Des noyaux accolés (a) et (b), sont séparés (f) : la fonction distance (b) vue 3D de la fonction distance (e) les érodés ultimes (d).

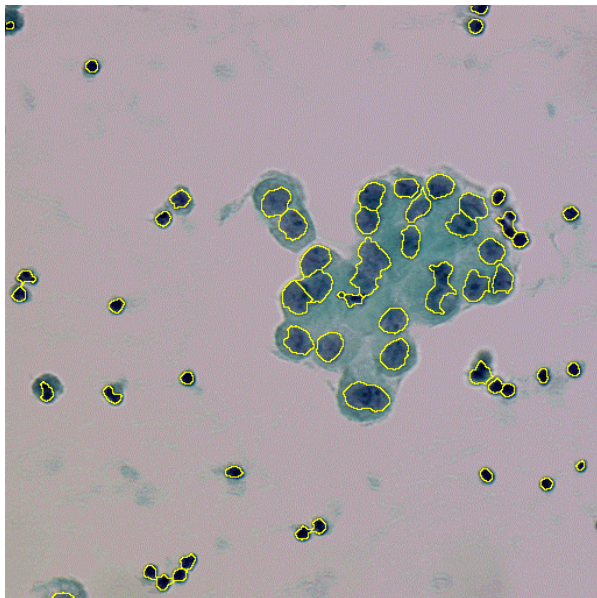


Figure 34.a. Une papille où les noyaux ont été séparés. Seuls les noyaux pouvant être séparés l'ont été.

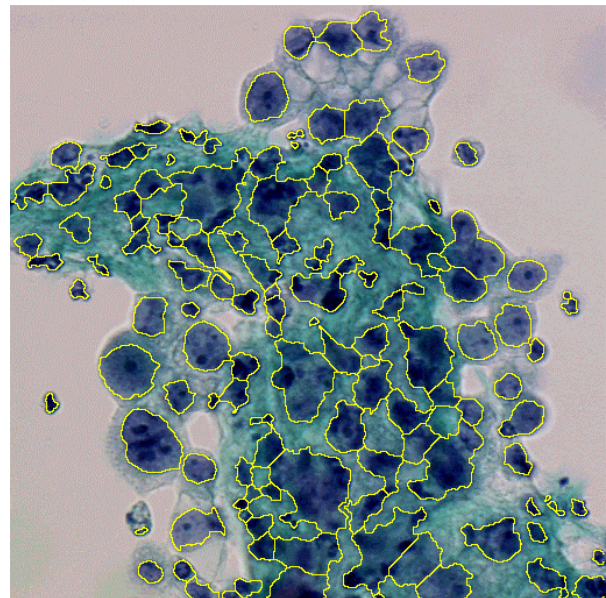


Figure 34.b. Une image d'amas où les noyaux ont été séparés.

Figure 34. Deux exemples de séparation de noyaux.

2.3.2.8. Division des cytoplasmes

A ce stade de notre stratégie de segmentation, nous avons des cytoplasmes qui contiennent soit un noyau, soit un amas de noyaux, soit plusieurs noyaux, soit plusieurs amas de noyaux. Nous désirons à travers cette ultime étape diviser tous les cytoplasmes afin qu'ils ne contiennent qu'un seul noyau ou amas de noyaux. Dans la majorité des cas, les lignes de séparation entre les cellules sont visibles et définissent les cytoplasmes se touchant. Mais dans

certaines configurations spatiales des cellules, ces transitions ne sont pas clairement définies et la seule information a priori qui puisse permettre de définir différentes zones séparées dans un cytoplasme réside dans les positions des noyaux à l'intérieur de ce même cytoplasme. Notre méthode doit donc tenir compte de ces deux informations : les transitions dans le vert entre les cytoplasmes et les noyaux à l'intérieur d'un cytoplasme.

Nous utilisons une méthode similaire à celle présentée au § 2.3.2.7. Nous devons donc disposer d'une image de marqueurs et d'une image représentant une fonction d'agrégation. Les marqueurs sont les noyaux précédemment obtenus. Cela permet de prendre en compte les positions respectives de ces derniers dans le cytoplasme à diviser (Figure 35 a et b). La fonction d'agrégation est la composante bleue lissée par un lissage exponentiel. Cette image respecte la forme du cytoplasme et les transitions internes au cytoplasme (ce qui satisfait notre premier critère). Les régions finales correspondant aux différents cytoplasmes sont exprimées par : $I_{\text{CytoplasmesSéparés}} = I_{CS} = WTS(Exp(I_{RGB}^B), I_{NS})$.

Ce sont des cytoplasmes contenant soit un noyau soit un amas de noyaux (Figure 35.c et d).

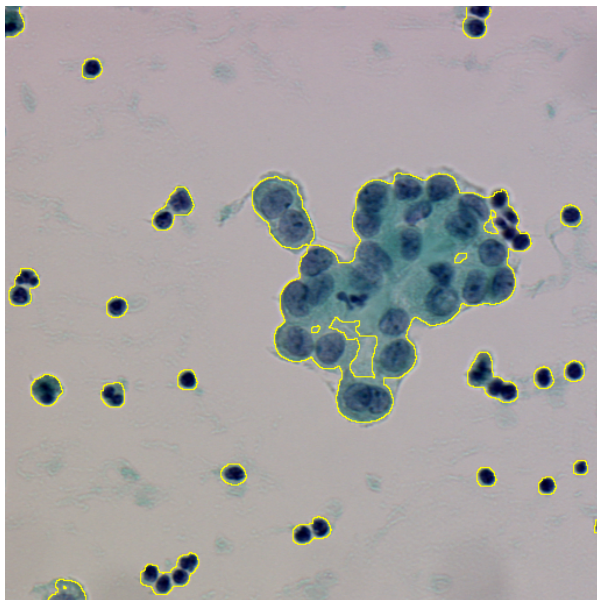


Figure 35.a. Image de papille avant séparation.

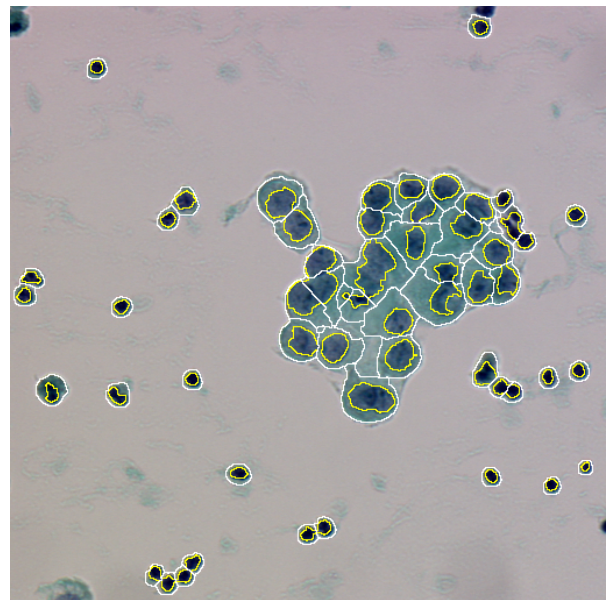


Figure 35.b. Image de papille après séparation.

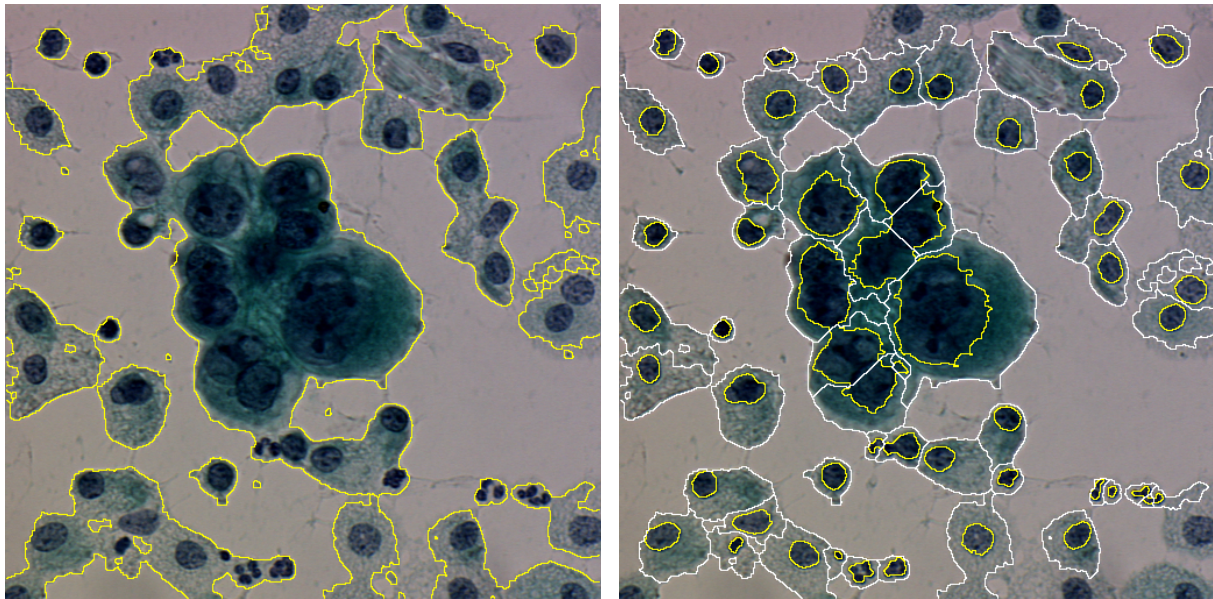


Figure 35.c. Image d'amas avant séparation.

Figure 35.d. Image d'amas après séparation.

Figure 35. Images présentant la séparation des cytoplasmes.

2.3.3. Nœuds dans des panneaux de de bois

Les images que nous avons à segmenter sont des images de planches de bois sur lesquelles sont présents des nœuds. Un nœud de bois est défini comme étant une zone de bois dure où se rejoignent une branche et le tronc d'un arbre. Ceux-ci doivent être extraits de façon automatique afin de réaliser une inspection des morceaux de bois [Kauppinen 95, Silven 96]. Les différents types de nœuds de bois sont les suivants :

1. Nœud creux,
 2. Nœud aride,
 3. Nœud recouvert,
 4. Nœud pourrissant,
 5. Nœud en feuilles,
 6. Nœud de bord.
- Nœud creux : nœud qui a grandi fermement à l'intérieur du bois qui l'entoure. La couleur de ce nœud peut, en périphérie, être très proche de celle du bois.
 - Nœud aride : nœud qui n'a pas pris complètement part au processus de développement du tronc de l'arbre. La couleur de ce nœud est habituellement plus foncée que la couleur d'un nœud creux et un anneau de couleur plus foncée entoure le nœud.
 - Nœud recouvert : nœud recouvert totalement ou partiellement par une éraflure en anneau. Comparativement à un nœud aride, l'anneau autour du nœud est plus épais.
 - Nœud pourrissant : nœud présentant des traces de pourrissement du bois. Visuellement, ce type de nœud n'est pas difficile à reconnaître car la texture du nœud est détériorée.
 - Nœud en feuilles : nœud qui s'est divisé et qui ne touche pas le bord du bois.

- Nœud de bord : nœud qui touche le bord.

Ici on ne s'attache pas à reconnaître les nœuds, mais seulement à les segmenter par notre méthode de segmentation couleur. Nous détaillons les différentes étapes de la stratégie d'optimisation de la fonction d'agrégation.

1. Pour choisir l'espace couleur le mieux adapté à notre méthode, nous avons segmenté manuellement 12 images couleur représentatives des nœuds pouvant être rencontrés. Deux images de chacune des classes ont été utilisées. L'espace *HSL* s'est avéré être le plus adapté (Tableau 7). Les Figure 36 (a) à (l) donnent les différentes images qui ont été utilisées dans cette étape.

Image	Type de nœud	RVB	XYZ	$L^*u^*v^*$	$L^*a^*b^*$	HSL	Chassery	Nbre de Points
1	nœud creux	52076	55556	29185	30865	55908	37856	592
2	nœud creux	21590	22745	11677	14392	23241	11217	378
3	nœud aride	10520	10682	6998	8900	10645	7627	80
4	nœud aride	64024	63596	46578	47540	63605	59465	500
5	nœud recouvert	14230	14181	13593	14064	14496	13390	123
6	nœud recouvert	59507	59137	47753	47127	59507	53646	492
7	nœud pourrissant	35691	35938	31526	29019	38354	19059	775
8	nœud pourrissant	26580	27649	28435	22468	28492	14882	480
9	nœud en feuilles	41591	41448	30180	32905	42453	25664	402
10	nœud en feuilles	18405	18469	20770	22016	18748	18839	280
11	nœud de bord	51120	49686	49014	48882	60361	53574	651
12	nœud de bord	41402	42174	30355	31538	41574	36754	576
Moyenne	Tous types	36395	36772	28839	29143	38115	29331	444

Tableau 7. Les contrastes pour plusieurs images de nœuds.

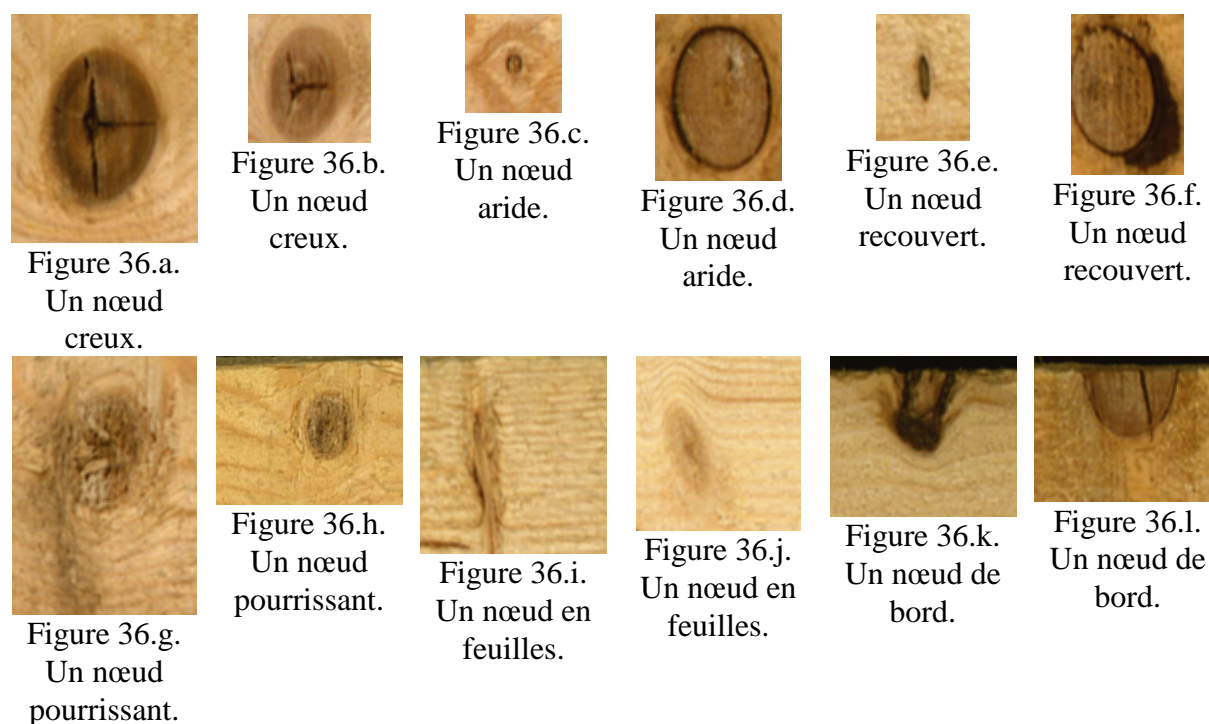


Figure 36. Les différents types de nœuds.

2. La métrique est une métrique euclidienne comme pour les autres applications.
3. Le calcul du Gradient est effectué dans l'espace HSL sélectionné en 1.
4. Pour choisir les marqueurs, on voudrait intuitivement extraire les zones marron d'une certaine gamme caractérisant les nœuds. On se trouve confronté à deux problèmes : le fond a la même couleur que les objets à extraire et suivant les nœuds leur couleur varie sur une très grande échelle : de très clair (Figure 36.h) à très foncé (Figure 36.d). La seule véritable information a priori dont on dispose réside dans le fait qu'un nœud est plus foncé que son entourage. Deux notions interviennent ici : la luminance et la saturation qui caractérisent l'un le fait que les nœuds sont plus foncés que le bois et l'autre le fait qu'un nœud présente une couleur marron plus saturée que le bois. Afin de prendre en compte ces deux informations nous utilisons, pour trouver les marqueurs, l'image I_{S_3} :

$$I_{S_3} = \frac{I_{HSL}^L + I_{HSL}^S}{2}$$

Cette image est seuillée en utilisant la méthode de Fisher (Figure 37.b). qui minimise la somme des variances à l'intérieur de chaque classe. Les trous présents dans les marqueurs sont comblés (Figure 37.c) et la zone d'influence est étendue (Figure 37.d).

$$I_{\text{MarquersNoeuds}} = I_{MNO} = E \left(s \left(I_{S_3} \right) \right)$$

5. Les nœuds (Figure 37.a) sont obtenus par la ligne de partage des eaux couleur (Figure 37.e) et l'image finale s'exprime par :

$$I_{\text{Noeuds}} = CWTS \left(\nabla (I_{HSL}), LAB(I_{MNO}), I_{HSL} \right)$$

Le paramètre α est fixé à 0,7. Ce paramètre a été déterminé à la suite à d'évaluations visuelles sur les images segmentées.



Figure 37.a.
L'image
originale.



Figure 37.b.
Les marqueurs
par la
méthode de
Fisher.



Figure 37.c.
Les
marqueurs
après
élimination
des trous.



Figure 37.d.
Les marqueurs
après
étiquetage et
érosion.



Figure 37.e.
Le nœud
segmenté.

Figure 37. Les différentes images obtenues lors de l'extraction des marqueurs et de la LPE.

La stratégie complète d'extraction des noyaux est exprimée sur la Figure 38 par un graphe d'opérateurs.

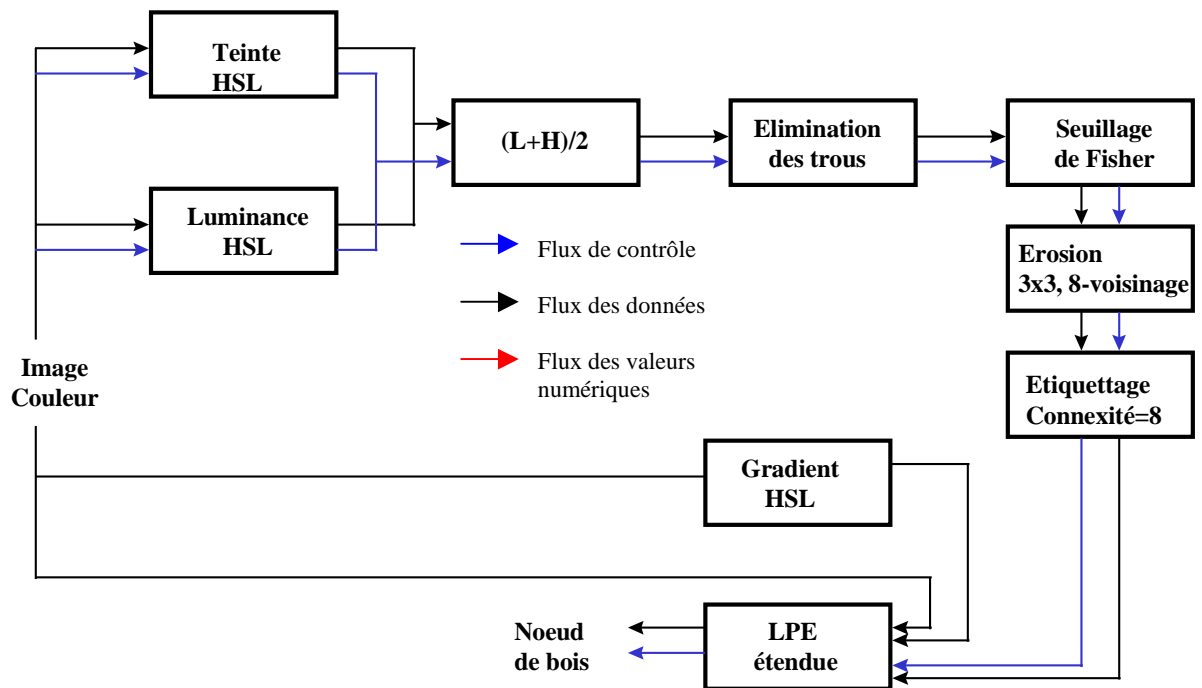
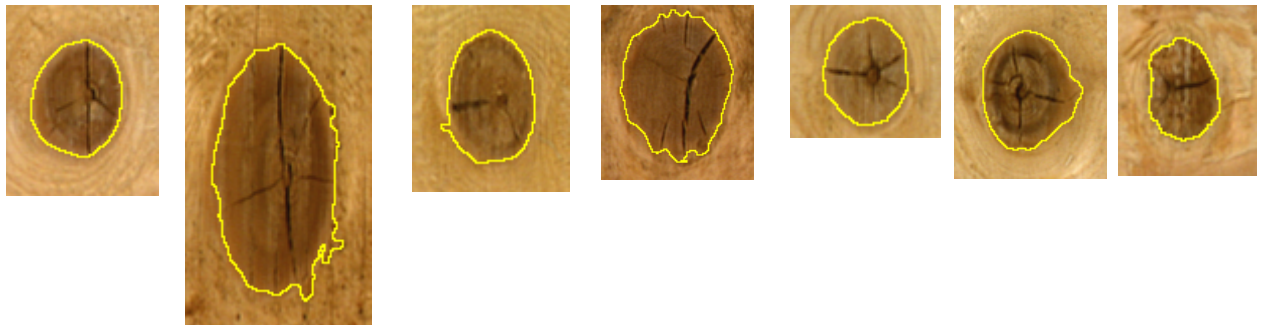


Figure 38. Graphe de la segmentation des nœuds de bois.

Voici quelques exemples des différentes classes de nœuds segmentés (Figure 39).

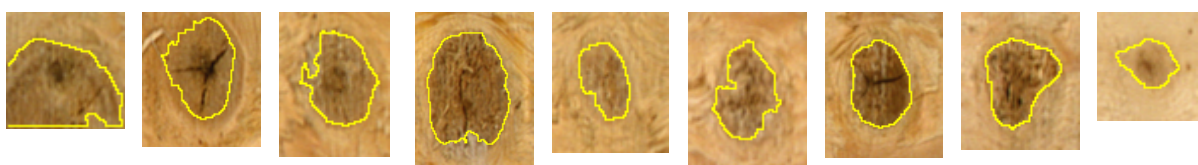
Nœuds creux



Nœuds arides



Nœuds pourrissants



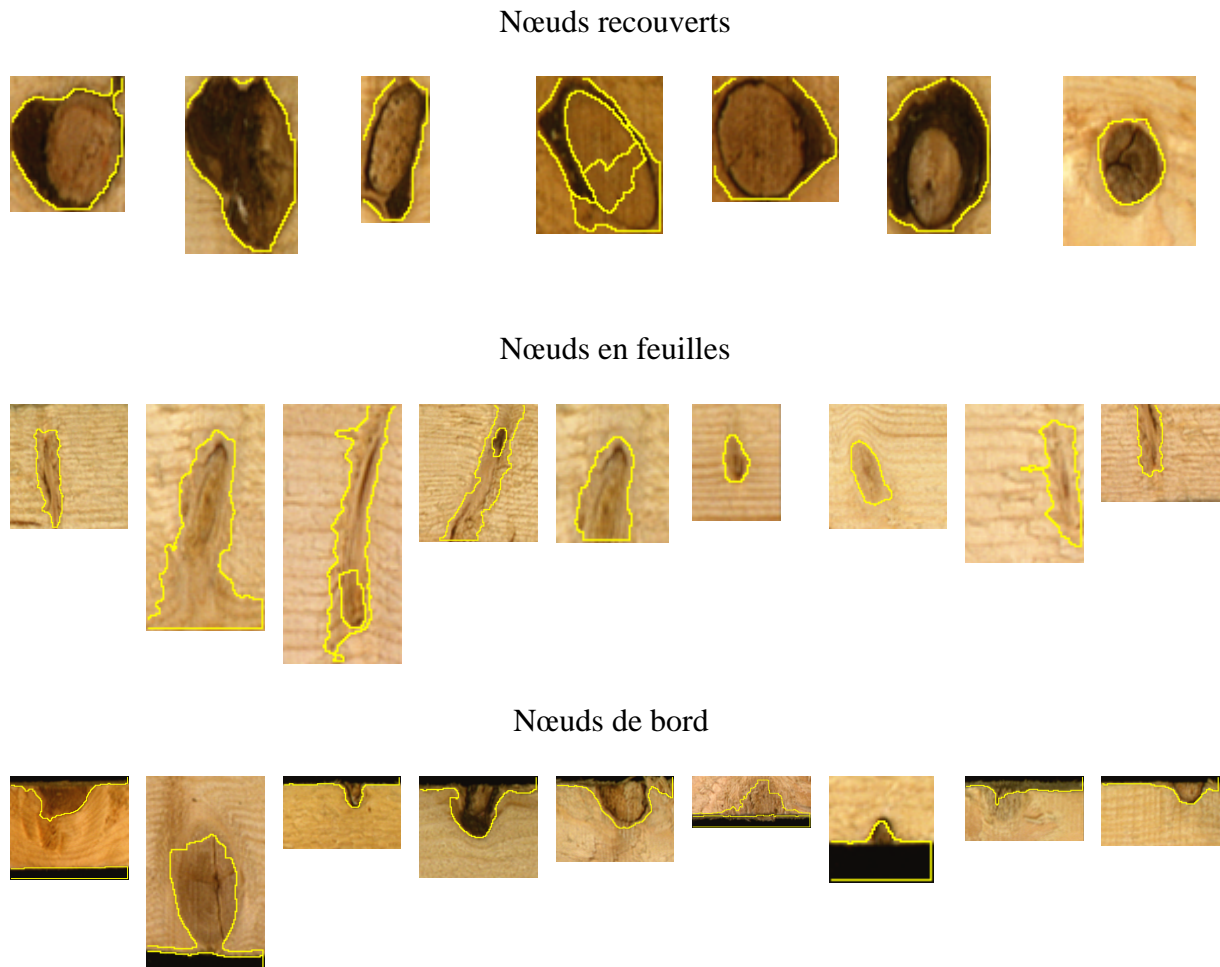


Figure 39. Exemples de nœuds segmentés.

On remarque que certains nœuds ont été segmentés en plusieurs parties et une fusion systématique des régions obtenues sera nécessaire pour achever la segmentation complète des nœuds. L'obtention de plusieurs régions provient de l'extension de la zone d'influence des marqueurs par érosion qui peut séparer en plusieurs marqueurs un marqueur qui était unique.

2.3.4. Conclusion sur les expérimentations

La méthode de segmentation présentée ici ne doit pas être vue comme une méthode générale de segmentation d'images couleur. Cette méthode est extrêmement liée au domaine des images et des informations a priori qui s'y rattachent et principalement la couleur. De plus la méthodologie d'utilisation nécessite de disposer de ces informations a priori et de déterminer complètement le problème à résoudre. Les objets doivent être connus et doivent être utilisés pour déterminer quel espace couleur utiliser. Comme nous l'avons vu, ce choix dépend de deux facteurs : la méthode utilisée et l'information recherchée dans les images.

Ainsi notre méthode de segmentation ne peut pas s'appliquer à tout type d'image couleur car elle nécessite un apport préalable d'informations a priori sur les objets. Néanmoins il semble difficile qu'une méthode de segmentation couleur puisse être universelle et permette d'extraire exactement tous les objets présents dans une image sans tenir compte d'informations a priori les concernant. C'est dans cette optique que doit être utilisée notre méthode de segmentation d'images couleur.

3. Caractérisation couleur

3.1. Introduction

Dans un système de vision par ordinateur, lorsque l'on se trouve confronté à un problème de reconnaissance d'objets, la première étape consiste en leur segmentation. L'étape suivante vise à extraire des informations de l'image. Par extraction d'informations on entend : fournir une description d'un objet à partir de sa modélisation dans l'image. Idéalement, un objet extrait lors de la segmentation est décrit simplement par une région : un ensemble de pixels connexes auxquels sont affectés une étiquette particulière et unique. La caractérisation consiste à associer à cet ensemble de pixels une description de l'objet par le biais de mesures quantitatives et/ou symboliques. Ici nous nous restreignons aux mesures quantitatives. A l'issue de la caractérisation on dispose, pour chaque objet, d'une description unique constituée d'un certain nombre de descripteurs ou d'attributs.

Certains attributs sont généraux et d'autres sont spécifiques à un certain type d'application à savoir la microscopie. Trois types de descripteurs sont principalement utilisés pour décrire un objet : la taille, la forme et la texture.

Dans ce paragraphe, nous allons définir différents attributs (dont certains sont classiques). Nous nous plaçons dans le cadre d'images couleur.

Les notations suivantes sont utilisées : $I_{c_1c_2c_3}$ désigne l'image couleur originale, I_S l'image des régions segmentée, o le label de la région à caractériser, $p(x,y)$ un point de l'image, $I_S(p)$ la valeur du pixel p dans l'image I_S , $Vois^B(p)$ l'ensemble des pixels voisins de p dans le voisinage B .

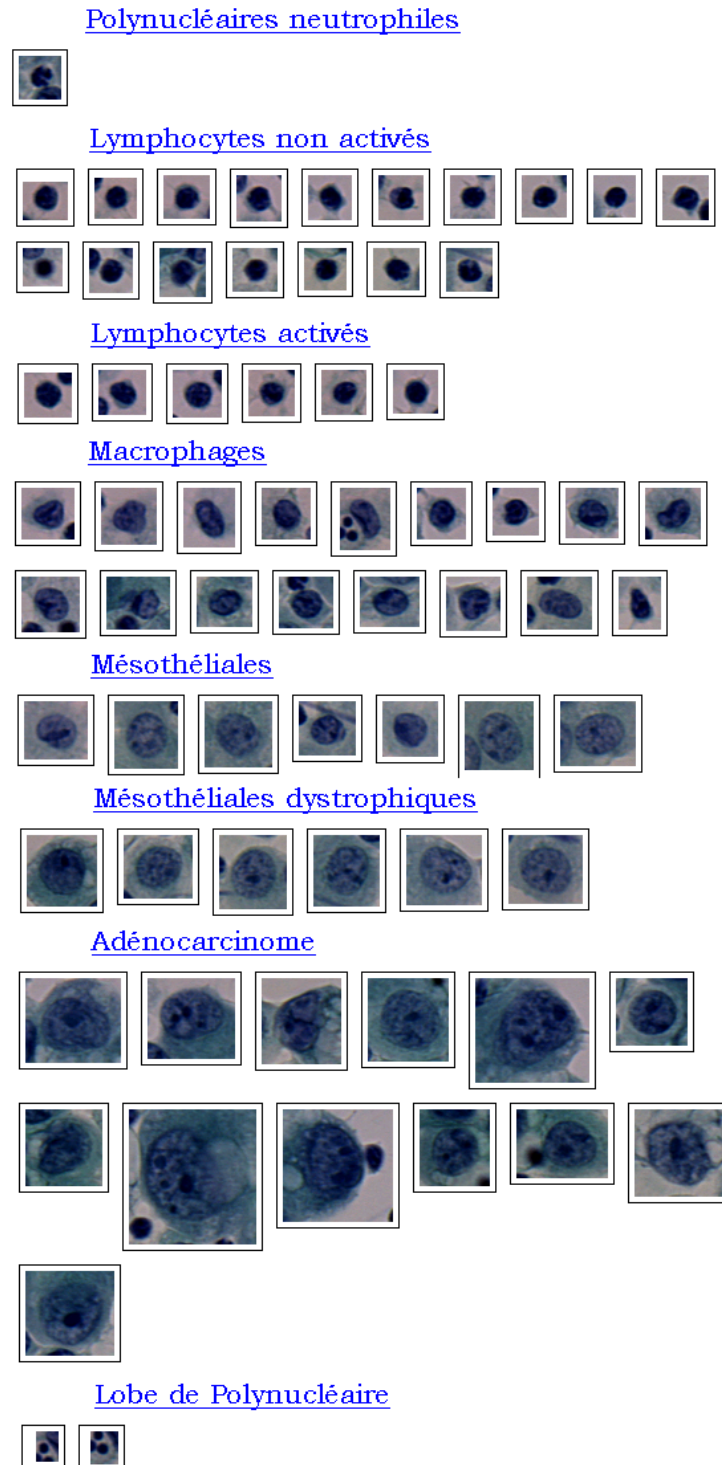


Figure 40. Des cellules de la cytologie des séreuses suivant leur classe d'appartenance.

3.2. Taille

Les descripteurs de taille sont les plus couramment utilisés. Ceci se comprend aisément étant donné qu'une distinction visuelle rapide peut permettre à un expert de décrire un objet par des termes tels que « grand » ou « petit » (à une échelle donnée). On définit ainsi la surface et le périmètre d'un objet o :

$$Surface(o) = Card\{p \in I_s / I_s(p) = o\}$$

$$Périmètre(o) = Card\{p \in I_s, I_s(p) = o / \exists m \in Vois^B(p) \text{ et } I_s(m) \neq o\}$$

Pour analyser la pertinence de ces deux paramètres, nous avons calculé la surface et le périmètre des cellules de la Figure 40. Une représentation graphique en est donnée ci dessous (Figure 41).

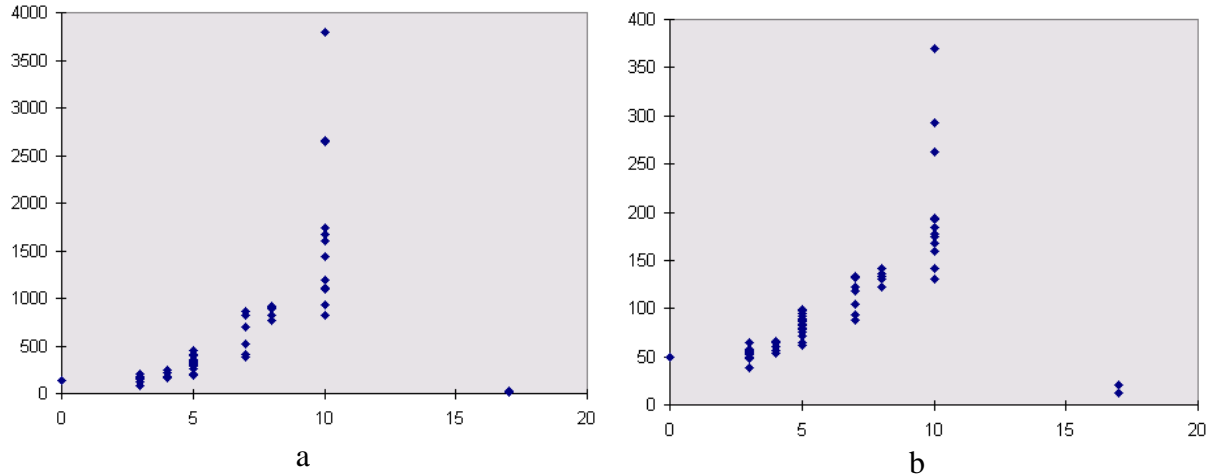


Figure 41 : caractérisation de cellules des séreuses par la surface (a) et par le périmètre (b). Les séries de noyaux correspondent aux séries verticales de points.

On constate que les attributs de surface et de périmètre ne suffisent pas à séparer les classes entre elles d'une part et que d'autre part ces deux attributs sont fortement corrélés entre eux (coefficient de corrélation de 0,98).

3.3. Forme

Les descripteurs de forme sont également très importants, il permettent d'avoir une description géométrique des objets à caractériser. Nous donnons ici la définition de la compacité, de l'élongation, de l'excentricité, de l'orientation, de la convexité et de la solidité d'un objet o .

$$Compacité(o) = \frac{4\pi * Surface(o)}{Périmètre(o)}$$

$$Elongation(o) = \frac{\sqrt{|M_x - M_y| + 4M_{xy}}}{Surface(o)}$$

$$Orientation(o) = \arctan\left(\frac{2M_{xy}}{M_x - M_y}\right)$$

$$Excentricité(o) = \frac{\left(\frac{M_x + M_y + \left(4M_{xy}^2 + (M_x - M_y)^2\right)}{2}\right) + \left(\frac{M_x + M_y - \left(4M_{xy}^2 + (M_x - M_y)^2\right)}{2}\right)}{(Surface(o))^2}$$

$$M_x = \frac{1}{Surface(o)} \sum_{i=1}^{Surface(o)} (x_i^2 - x_G^2) \text{ et } M_y = \frac{1}{Surface(o)} \sum_{i=1}^{Surface(o)} (y_i^2 - y_G^2)$$

$$M_{xy} = \frac{1}{Surface(o)} \sum_{i=1}^{Surface(o)} (x_i y_i - x_G y_G)$$

avec $p_i(x_i, y_i)$, $I_S(p) = o$ et $G(x_G, y_G)$ le centre de gravité de l'objet o .

$$Convexité(o) = \frac{PérimètreConvexe(o)}{Périmètre(o)}$$

$$Solidité(o) = \frac{Surface(o)}{SurfaceConvexe(o)}$$

Le *PérimètreConvexe* et la *SurfaceConvexe* utilisent l'enveloppe convexe d'un objet. Celle-ci est obtenue grâce à la méthode des tangentes.

Afin de rendre compte des différences entre ces paramètres nous les avons calculés sur différentes formes numérotées de 1 à 9 (Figure 42.a). Leur enveloppe convexe est donnée par la Figure 42.b.

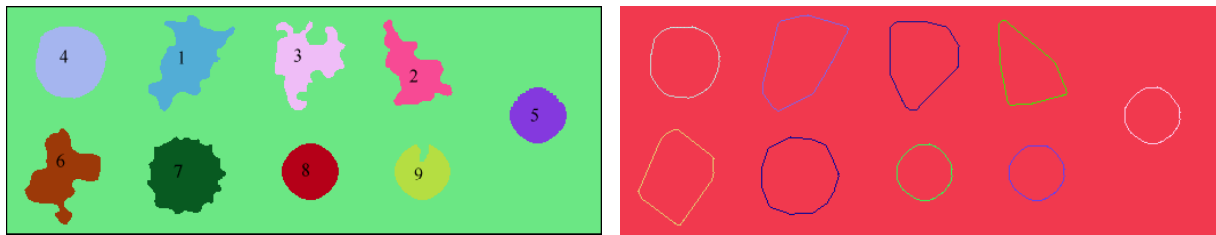
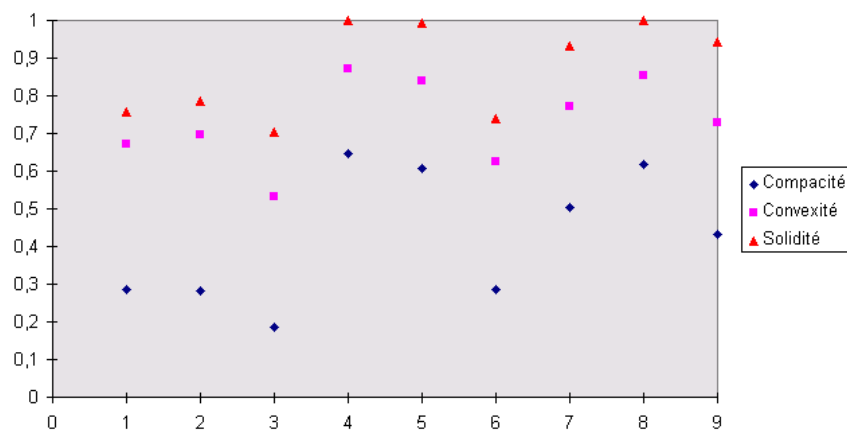


Figure 42. Des objets de différentes formes (a) et leurs enveloppes convexes respectives (b).

Les résultats sont présentés sur la Figure 43 (de a à e).



a : compacité, convexité et solidité.

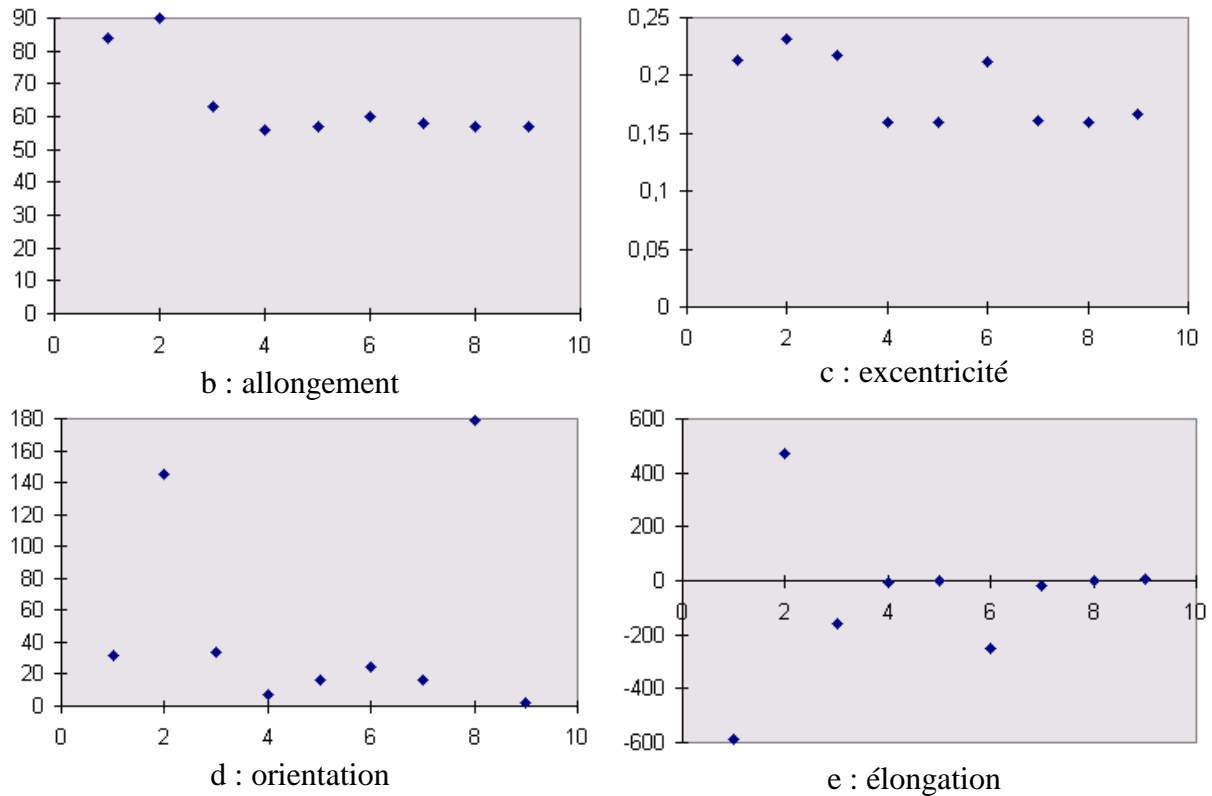


Figure 43. Caractérisation des formes de la Figure 42.a.

On constate que la compacité, la convexité et la solidité se comportent sensiblement de la même manière et permettent d'avoir une estimation simple de la forme d'un objet. L'allongement fait ressortir les formes 1 et 2 comme plus allongées, mais les formes 3 et 6 qui pourtant sont allongées ne ressortent pas par rapport aux autres. Par contre l'élongation permet de distinguer les formes 3 et 6. L'excentricité donne des résultats identiques, les formes « bizarres » que sont les formes 1, 2, 3 et 6 sont bien distinctes des autres formes. L'orientation ne donne pas de résultats satisfaisants : elle ne peut pas être stable pour des objets ronds car elle est trop sensible aux faibles variations. L'interprétation des différents descripteurs de forme est difficile, l'évaluation de leur apport d'information (au sens de la discrimination entre les différentes formes) est loin d'être triviale. Car bien qu'un attribut ne soit pas fortement discriminatoire au premier ordre, celui-ci peut s'avérer discriminatoire au second ordre lorsqu'il est combiné avec un autre paramètre.

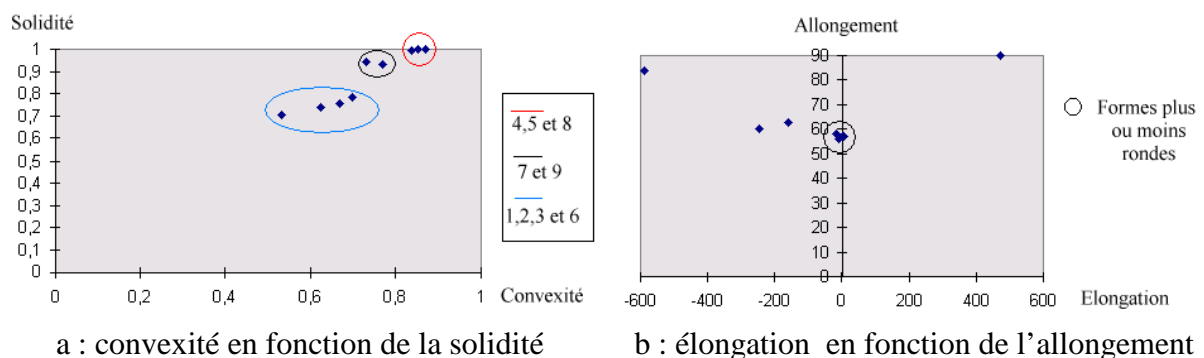


Figure 44. Le pouvoir discriminatoire d'un paramètre en fonction d'un autre.

On peut en effet constater sur la Figure 44 que deux paramètres corrélés entre eux (solidité et convexité) ou bien deux paramètres non corrélés mais dont l'un semble peu informant (l'allongement) peuvent permettre de discriminer différentes formes.

3.4. Texture couleur

La texture est une caractéristique importante pour l'analyse d'un grand nombre d'objets. Cela peut se vérifier sur des images provenant de contextes différents. En dépit de l'importance et de l'ubiquité de la texture pour l'interprétation des données d'une image (couleur ou niveaux de gris), une approche formelle ou une définition précise de la texture n'existe pas. Les techniques de discrimination de la texture sont, pour la plupart, obtenues *ad hoc* et dépendent des objets à caractériser.

Ce dernier point accentue la part de subjectivité contenue dans le processus visuel d'évaluation d'une texture. Cette évaluation pose plusieurs problèmes majeurs :

1. Il existe différentes approches pour définir la notion même de texture, chacune restreignant sa signification. La méthode employée reste donc comme une façon déterminante d'évaluer une texture.
2. L'évaluation d'une texture dépend du niveau d'observation, suivant les paramètres utilisés dans la méthode de caractérisation d'une texture, des textures fines peuvent être considérées comme plus grossières à un niveau d'analyse différent.
3. Les évaluations des experts du domaine sont souvent très différentes car subjectives et les limites des différentes classes de texture ne sont pas souvent très bien définies.

Au vu de ces remarques, le choix d'une méthode de caractérisation de la texture ne peut se faire qu'en connaissance du contexte des objets à reconnaître et après une phase de concertation avec les experts du domaine afin de « modéliser » dans un premier temps le concept même de la texture attachée aux objets à caractériser. Ce problème est évidemment capital pour l'obtention de données satisfaisantes caractérisant une texture.

Avant de préciser exactement les approches que nous avons choisies d'utiliser pour caractériser une texture, voyons brièvement quelques unes des techniques majeures employées en reconnaissance de texture (au sens de leur caractérisation). Cependant nous ne traitons que les techniques employées en microscopie.

3.4.1. Méthodes usuelles

Les méthodes employées pour caractériser une texture sont très diverses [Haralick 79]. Les plus communes sont :

- Les mesures statistiques dont la matrice de cooccurrence [Walker 94, Jorensen 96, Angot 98]. Elles permettent de disposer de mesures statistiques sur les objets. Cela permet de mesurer la distribution et la dépendance spatiale des niveaux de gris.
- Les caractéristiques géométriques statistiques [Walker 96]. Elles décomposent une texture donnée en niveaux de gris par un ensemble d'images seuillées. Certaines propriétés géométriques des régions sont mesurées pour chacune des images seuillées et des mesures statistiques basées sur ces propriétés géométriques sont calculées.
- Les opérateurs de Morphologie Mathématique [Jackway 95b, Wolf 94, Giménez-Mas 94, Moragas 93, Garcia-Bonafé 96]. Ils permettent de diviser la texture en

plusieurs régions. Les caractéristiques extraites portent alors sur l'organisation de ces régions et sur leur position relative. Citons quelques-uns de ces opérateurs : chapeau haut de forme, érosion, dilatation, ligne de partage des eaux.

- L'analyse Markovienne [Dawson 93]. Elle est basée sur les probabilités de transition entre les niveaux de gris.

Le choix d'une méthode pour caractériser une texture est un problème délicat. Nous avons choisi d'utiliser des méthodes fondées sur des propriétés morphologiques. Cependant, comme nous l'avons vu, l'étude d'une texture dépend du niveau d'analyse auquel on se place. Pour cette raison, nous avons choisi d'utiliser trois méthodes qui opèrent à des niveaux de précision différents, permettant ainsi d'avoir un certain point de vue relativement « multi-échelle » de la texture. Les caractéristiques que nous avons utilisées sont toutes obtenues à partir d'images couleur. En effet, la couleur, si elle est importante pour la segmentation d'une image, l'est également pour la caractérisation d'une texture (car caractériser une texture par des mesures morphologiques revient à la diviser en régions et donc à la segmenter sous un certain point de vue). Les paragraphes qui suivent présentent les méthodes que nous avons employées. Elles sont illustrées par des résultats sur des images de cellules.

3.4.2. Mesures Statistiques

On utilise des mesures classiques communément utilisées en caractérisation : moyenne, variance et écart-type. Cependant nous calculons la moyenne et la variance pour chacune des composantes couleur mais également la variance et l'écart-type pour la couleur. Pour un objet o , dans l'espace couleur $C_1C_2C_3$, avec $C_i \in \{C_1, C_2, C_3\}$ et un point tel que $I_S(p)=o$ on définit :

$$\begin{aligned} \text{Moyenne}_{C_i}(o) &= \frac{\sum I_{C_1C_2C_3}^{C_i}(P)}{\text{Surface}(o)} & \text{Variance}_{C_i}(o) &= \frac{\sum (I_{C_1C_2C_3}^{C_i}(P))^2 - \frac{(\sum I_{C_1C_2C_3}^{C_i}(P))^2}{\text{Surface}(o)}}{\text{Surface}(o)} \\ \sigma_{C_i}(o) &= \sqrt{\text{Variance}_{C_i}(o)} & \text{VarianceCouleur}(o) &= \frac{\sum_{k=1}^3 \text{Variance}_{C_k}(o)}{3} \\ \sigma(o) &= \frac{\sum_{k=1}^3 \sigma_{C_k}(o)}{3} \end{aligned}$$

3.4.3. Ligne de partage des eaux multi-échelle

L'approche multi-échelle a été introduite par Paul T. Jackway [Jackway 95, 95b] à travers la ligne de partage des eaux. L'approche proposée a l'avantage d'utiliser directement les minima présents dans une image. Puisque les lignes de partage des eaux correspondent chacune à des minima régionaux de l'image sur laquelle elles sont calculées, ce nombre de minima peut être réduit (et donc par conséquent le nombre de régions obtenues) en réduisant le nombre de minima régionaux de l'image (par une modification homotopique). Cette approche est l'approche classique de la ligne de partage des eaux [Vincent 90].

L'approche multi-échelle est une façon de contrôler la segmentation par une simplification de la surface sur laquelle elle est réalisée. La théorie du filtrage multi-échelle [Witkin 83] repose sur une simplification de l'image originale grâce à l'utilisation d'un

paramètre d'échelle qui donne un effet de segmentation multi-échelle. L'effet de ce type de segmentation réside dans la décroissance monotone du nombre de régions segmentées lorsque le paramètre d'échelle augmente [Jackway 94]. L'approche multi-échelle permet donc de décrire une image suivant plusieurs échelles en augmentant les relations entre les pixels grâce à une simplification. La propriété de décroissance monotone s'exprime par le fait que si un signal est présent à une échelle donnée, il doit être présent à des échelles plus faibles.

Pour assurer cette propriété dans le cadre de la ligne de partage des eaux, il faut s'assurer que le nombre de régions obtenues à différentes échelles de simplifications décroît de façon monotone lorsque l'échelle augmente. Ceci est réalisé grâce à la modification homotopique de l'image de gradient par reconstruction morphologique [Vincent 90]. Voyons ceci plus en détail.

Considérons une image I dont le domaine de définition est $D_I \subset \mathbb{R}^2$. I étant supposée ne prendre que des valeurs discrètes dans l'intervalle $[0, N]$, N étant un entier positif arbitraire.

$$I \left(\begin{array}{l} D_I \subset \mathbb{R}^2 \rightarrow \{0, 1, \dots, N\} \\ p \rightarrow I(p) \end{array} \right)$$

Définissons maintenant I^M comme la reconstruction morphologique de I où γ est la fonction de marquage utilisée pour sélectionner les minima à être retenus (d'après la définition donnée dans [Vincent 93])

$$I^M(p) = \rho_I(\gamma)(p)$$

En appliquant cette théorie à la ligne de partage des eaux, il est simple d'imposer des minima régionaux (et par conséquent des bassins versants) seulement aux endroits où se situent des minima régionaux de l'image lissée. Soit g_σ une fonction de lissage où σ représente l'échelle du lissage, on définit alors :

$$R_{\min}(\rho_{|\nabla(I \otimes g_\sigma)}(\gamma)) = R_{\min}(I \otimes g_\sigma) \text{ pour tout } \sigma \geq 0$$

en utilisant la fonction de marquage,

$$\gamma(p) = \begin{cases} I(p) & \text{si } p = m_i / m_i \in M \\ \infty & \text{ailleurs} \end{cases}$$

M définissant l'ensemble des q minima régionaux de I

Un exemple de modification homotopique d'une fonction de gradient est donné par la Figure 45 où le gradient modifié ne conserve que les minima qui correspondent à des minima dans la fonction originale.

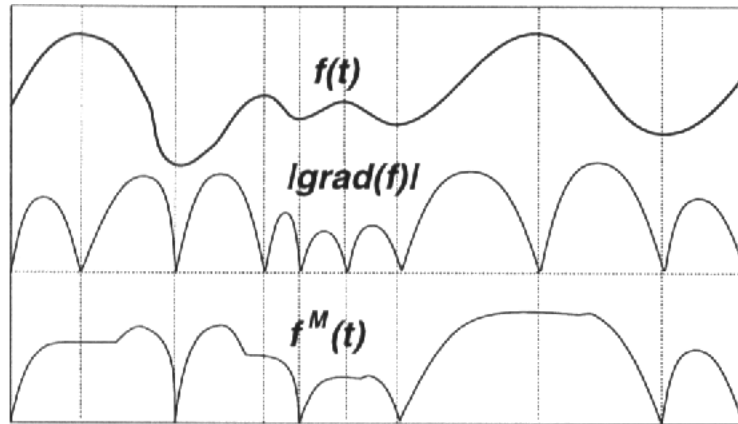


Figure 45. Modification homotopique du gradient d'une fonction d'après [Jackway 95b].

Le résultat majeur défini par Jackway établit le théorème de la monotonie multi-échelle pour les régions de la ligne de partage des eaux. Soient $\text{Card}[WTS(I)]$ le cardinal de l'ensemble des régions obtenues par la ligne de partage des eaux d'une image I , $|\nabla(I)|$ le gradient modifié de façon homotopique (voir ci-dessus) et $I \otimes g_\sigma$ le lissage de l'image I par une fonction g_σ , alors pour différentes échelles $0 < \sigma_1 < \sigma_2$ on a :

$$\text{Card}\left[WTS\left(|\nabla(I \otimes g_{\sigma_2})|\right)\right] \leq \text{Card}\left[WTS\left(|\nabla(I \otimes g_{\sigma_1})|\right)\right] \leq \text{Card}\left[WTS(|\nabla(I)|)\right]$$

L'application de cette méthode aux images couleur se fait naturellement : le seul changement survient dans la définition de la fonction de marquage,

$$\gamma(p) = \begin{cases} I_{C_1 C_2 C_3}(p) & \text{si } p = m_i / m_i \in M \\ \infty & \text{ailleurs} \end{cases}$$

M définissant l'ensemble des q minima régionaux de $I_{C_1 C_2 C_3}$

m étant un minima régional ssi $\nexists p / p \in \text{Vois}^\square(m)$ et avec $d(p, m) \leq 0$

Le gradient et la notion de distance restent définis comme dans le cadre de notre segmentation couleur. La fonction de lissage g_σ que nous avons utilisée est un lissage exponentiel. En appliquant la méthode à un objet couleur (ici une cellule) à différentes échelles, on obtient les résultats suivants dans l'espace RVB :

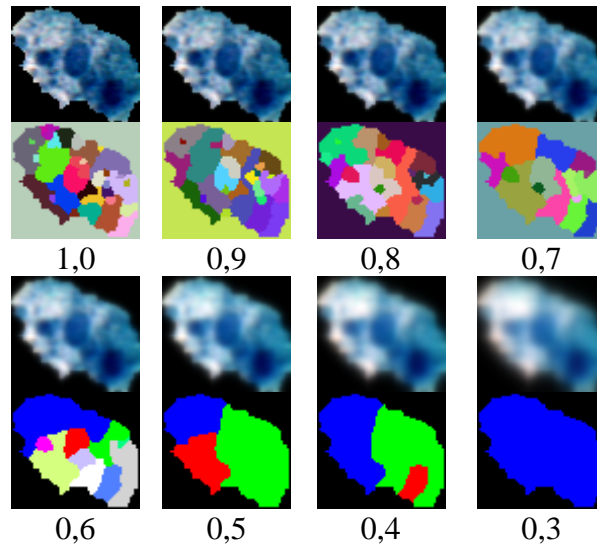


Figure 46. Les régions obtenues par la ligne de partage des eaux multi-échelle d'une cellule des séreuses.

Coefficient du lissage	1,0	0,9	0,8	0,7	0,6	0,5	0,4	0,3
Nombre de régions	33	27	23	15	9	3	3	1

Tableau 8. Le nombre de régions suivant le coefficient du lissage simulant l'aspect multi-échelle.

Les images, lissées par un filtre exponentiel, sont présentées de la manière suivante : en haut l'image couleur lissée et en bas, les régions obtenues. L'échelle du lissage est indiquée sous chaque paire d'image. On peut vérifier sur le Tableau 8 ainsi que sur la Figure 46 que le nombre de régions obtenues par la ligne de partage des eaux décroît lorsque le coefficient multi-échelle σ augmente, ceci confirme la propriété de monotonie de la ligne de partage des eaux multi-échelle.

Nous disposons donc d'une méthode de partitionnement d'un objet en régions, il nous reste à extraire des caractéristiques quantitatives de ces régions. Bien que l'on puisse extraire des caractéristiques à différents niveaux d'échelle, nous avons choisi de n'extraire des caractéristiques qu'à un seul niveau d'échelle. Le coefficient σ que nous avons retenu a pour valeur 0,7. Notons N le nombre de régions r_i associées à l'objet o à caractériser, on définit pour chaque région :

$$Taille_WTS(r_i, o) = Card\{P / I_s(P) = o \text{ et } P \in r_i\}$$

$$Couleur_WTS(r_i, o) = \sqrt{\left(\frac{\sum I_{C_1 C_2 C_3}^{C_1}(P)}{Taille_WTS(r_i, o)}\right)^2 + \left(\frac{\sum I_{C_1 C_2 C_3}^{C_2}(P)}{Taille_WTS(r_i, o)}\right)^2 + \left(\frac{\sum I_{C_1 C_2 C_3}^{C_3}(P)}{Taille_WTS(r_i, o)}\right)^2}$$

Ensuite on associe à chaque région r_i une valeur mesurant la distance de chaque région r_i aux frontières de l'objet : $Position_WTS(r_i, o)$. Cela permet de mesurer la position des

différentes régions à l'intérieur de l'objet o et de dégager non seulement sa structure mais également son organisation. Les caractéristiques associées à un objet o peuvent alors être définies :

$$\text{CouleurMoyenne_WTS}(o) = \text{CM_WTS}(o) = \frac{1}{N} \sum_{i=1}^N \text{Couleur_WTS}(r_i, o)$$

$$\text{TailleMoyenne_WTS}(o) = \text{TM_WTS}(o) = \frac{1}{N} \sum_{i=1}^N \text{Taille_WTS}(r_i, o)$$

$$\text{CouleurPositionMoyenne_WTS}(o) = \text{CPM_WTS}(o) = \frac{1}{N} \sum_{i=1}^N \text{Couleur}(r_i, o) * \text{Position_WTS}(r_i, o)$$

$$\sigma_{\text{TM_WTS}}(o) = \sqrt{\frac{1}{N} \sum_{i=1}^N (\text{Taille_WTS}(r_i, o) - \text{TM_WTS}(o))^2}$$

$$\sigma_{\text{CM_WTS}}(o) = \sqrt{\frac{1}{N} \sum_{i=1}^N (\text{Couleur_WTS}(r_i, o) - \text{CM_WTS}(o))^2}$$

$$\sigma_{\text{CPM_WTS}}(o) = \sqrt{\frac{1}{N} \sum_{i=1}^N (\text{Couleur_WTS}(r_i, o) * \text{Position_WTS}(r_i, o) - \text{CPM_WTS}(o))^2}$$

3.4.4. Textons

Définis par Julesz [Julesz 81], les textons sont les éléments de base sur lesquels la perception visuelle se pré-focalise. La texture est alors définie comme un ensemble de régions homogènes. Afin de définir ces textons, nous devons nous placer dans le cadre d'un modèle de texture hiérarchique établissant qu'une texture est composée d'un ensemble de primitives i.e. de textons [Julesz 81, Wolf 95]. Comme nous l'avons vu précédemment, l'étude d'une texture peut se faire à différentes échelles (modèle des lignes de partage des eaux multi-échelle), mais l'utilisation d'un modèle hiérarchique de texture suppose que chaque texton défini au niveau d'observation i est constitué d'éléments du niveau d'observation $i-1$. Le pixel représentant le plus petit niveau d'observation. Les textons sont définis comme un ensemble de pixels qui sont homogènes suivant un certain nombre de règles. Ce critère d'homogénéité est généralement basé sur la différence de niveaux de gris entre des pixels adjacents.

La méthode que nous avons utilisée pour extraire les textons d'un objet a été introduite par Wolf [Wolf 95]. Elle définit des règles divisant la texture en trois zones de texture : Montagne (M), Vallée (V) et Transition (T). Une montagne est entourée de régions inférieures, une vallée est entourée de régions d'altitudes supérieures, une pente fait la liaison entre les montagnes et les vallées. Suivant un modèle hiérarchique de texture, les régions sont constituées de pixels et l'affectation d'un pixel à l'une des trois classes de texture (M, V ou T) utilise les informations locales de voisinage. Wolf a défini des règles d'affectation qui permettent de situer un pixel dans l'une des trois classes. Ceci n'a été fait qu'en niveaux de gris et nous avons généralisé le modèle afin de l'appliquer aux images couleur.

La méthode peut s'apparenter à une méthode de croissance de régions et utilise les règles d'agrégation suivantes :

Soit I une image couleur, P un point de cette image ($P \in I_{C_1C_2C_3}$), $P_i \in \text{Vois}^\square(P)$ et $O(0,0,0)$ l'origine du repère utilisé dans l'espace couleur $C_1C_2C_3$.

Règle 1 :

Chaque pixel P vérifiant un ou deux critères est affecté à l'une des classes parmi M,V ou T.

$$(1) \left\| \overrightarrow{MP} \right\| > \min \left(\left\| \overrightarrow{MP_i} \right\| \right) + \delta$$

$$(2) \left\| \overrightarrow{OP} \right\| > \max \left(\left\| \overrightarrow{OP_i} \right\| \right) - \delta$$

$$P \rightarrow M \text{ si (1)}$$

$$P \rightarrow V \text{ si (2)}$$

$$P \rightarrow T \text{ si (1) et (2)}$$

avec δ une constante dont nous précisons la valeur par la suite.

Règle 2 :

Chaque pixel P qui n'a pas été affecté à l'une des classes M,V ou T, mais dont l'un des voisins est assigné, est affecté à M,V ou T suivant les conditions :

$$\text{Pour } P / \exists P_i \in \text{Vois}^\square(P), P_i \in \{M, V, T\}$$

$$P \rightarrow M \text{ si } \left\| \overrightarrow{OP} \right\| \geq \min \left(\left\| \overrightarrow{OP_i} \right\| \right) / P_i = M$$

$$P \rightarrow V \text{ si } \left\| \overrightarrow{OP} \right\| \leq \max \left(\left\| \overrightarrow{OP_i} \right\| \right) / P_i = V$$

$$P \rightarrow P \text{ si } \min \left(\left\| \overrightarrow{OP_i} \right\| \right) \leq \left\| \overrightarrow{OP} \right\| \leq \max \left(\left\| \overrightarrow{OP_i} \right\| \right) / P_i = T$$

On remarque cependant quelques cas particuliers peuvent apparaître à l'issue de cette seconde règle :

$$\text{si } P = M : \left\| \overrightarrow{OP_i} \right\| \leq \max \left(\left\| \overrightarrow{OP_i} \right\| \right) \text{ avec } P_i = M \text{ et}$$

$$\left\| \overrightarrow{OP_i} \right\| \leq \max \left(\left\| \overrightarrow{OP_i} \right\| \right) \text{ avec } P_i = T$$

$$\text{si } P = V : \left\| \overrightarrow{OP_i} \right\| \geq \min \left(\left\| \overrightarrow{OP_i} \right\| \right) \text{ avec } P_i = M \text{ et}$$

$$\left\| \overrightarrow{OP_i} \right\| \geq \min \left(\left\| \overrightarrow{OP_i} \right\| \right) \text{ avec } P_i = S$$

$$\text{si } P = T : \left\| \overrightarrow{OP_i} \right\| \leq \max \left(\left\| \overrightarrow{OP_i} \right\| \right) \text{ avec } P_i = V \text{ et}$$

$$\left\| \overrightarrow{OP_i} \right\| \geq \min \left(\left\| \overrightarrow{OP_i} \right\| \right) \text{ avec } P_i = M$$

Ces cas particuliers sont résolus par la règle 3.

Règle 3 :

Chaque paire de pixels adjacents P_1 et P_2 qui vérifient l'une des trois conditions suivantes sont affectés à la classe T.

$$(1) P_1 = M \text{ et } P_2 = V \quad \text{et} \quad \left\| \overrightarrow{OP_1} \right\| \leq \left\| \overrightarrow{OP_2} \right\|$$

$$(2) P_1 = M \text{ et } P_2 = T \quad \text{et} \quad \left\| \overrightarrow{OP_1} \right\| \leq \left\| \overrightarrow{OP_2} \right\|$$

$$(1) P_1 = V \text{ et } P_2 = T \quad \text{et} \quad \left\| \overrightarrow{OP_1} \right\| \geq \left\| \overrightarrow{OP_2} \right\|$$

Règle 4 :

Une région de pixels connexes qui n'appartient à aucune des classes M,V ou T, est affectée à la classe M ou V si cette région est entourée exclusivement de montagnes ou de vallées. Si cette région est entourée à la fois de montagnes et de vallées, cette région est affectée à la classe T.

Une fois ces règles régissant la croissance de région établies, l'algorithme général est :

```

tant que des pixels sont affectés
  Règle 1
  Règle 2
  Règle 3
fin tant que
Règle 4

```

Le paramètre δ est calculé en fonction de l'objet à caractériser. Il doit être suffisamment précis pour permettre une caractérisation optimale de la texture d'un objet. Cependant il n'est pas souhaitable que ce paramètre soit fixé de manière arbitraire et empirique. En effet un seuil fixe n'est pas souhaitable, car la texture dépend généralement de l'objet que l'on cherche à caractériser. Au vu de ces remarques, le seuil δ est fixé comme étant égal à l'écart-type de la couleur de l'objet à caractériser :

$$\delta = \frac{\sigma_{c_1}(o) + \sigma_{c_2}(o) + \sigma_{c_3}(o)}{3}$$

Voyons un exemple du déroulement de la croissance de régions sur un objet provenant de la cytologie des séreuses ; les classes Montagne, Transition et Vallée vont respectivement du plus foncé au plus clair (Figure 47).

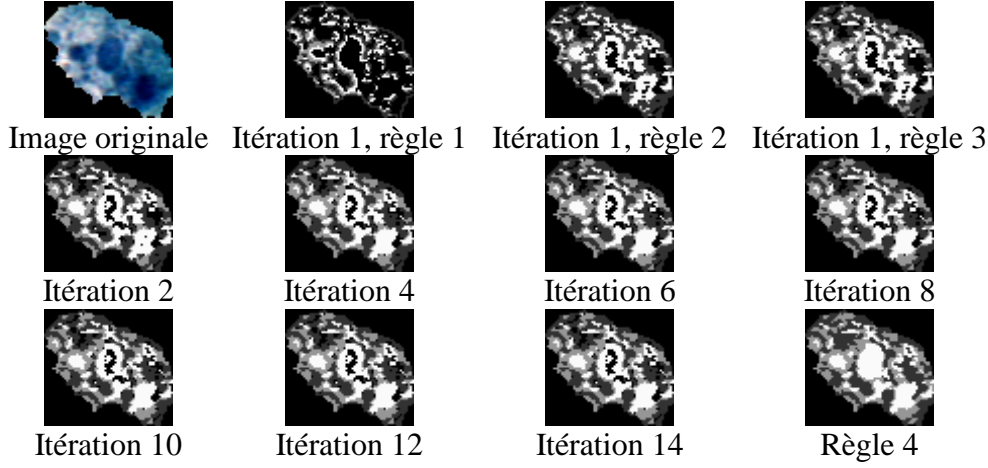


Figure 47. Les étapes de l'extraction de texture d'un objet dans l'espace *RVB*.

En comparant l'image finale des régions (ou textons) obtenues, on remarque que cela correspond très bien aux zones texturées présentes sur l'objet. La méthode permet donc une extraction fiable et précise des textons d'un objet. A partir de ces textons les caractéristiques calculées se rapprochent de celles calculées pour la méthode de ligne de partage des eaux multi-échelle. On définit pour un objet o , M_i comme les régions appartenant à la classe M , V_j comme le nombre de régions appartenant à la classe V ,

$$\text{Nombre_Montagnes}(o) = NM(o) = \text{Card}\{M_i\}$$

$$\text{Nombre_Vallées}(o) = NV(o) = \text{Card}\{V_j\}$$

$$\text{Taille}(C, o) = \text{Card}\{P / I_s(P) = o \text{ et } P \in C\} \text{ avec } C \in \{M, V\}$$

$$\text{TailleMoyenne_Montagnes}(o) = TMM(o) = \frac{\sum_{i=1}^{NM(o)} \text{Taille}(M_i, o)}{NM(o)}$$

$$\text{TailleMoyenne_Vallées}(o) = TMV(o) = \frac{\sum_{j=1}^{NV(o)} \text{Taille}(V_j, o)}{NV(o)}$$

$$\text{Couleur}(C, o) = \sqrt{\left(\frac{\sum I_{C_1 C_2 C_3}^{C_1}(P)}{\text{Taille}(C, o)}\right)^2 + \left(\frac{\sum I_{C_1 C_2 C_3}^{C_2}(P)}{\text{Taille}(C, o)}\right)^2 + \left(\frac{\sum I_{C_1 C_2 C_3}^{C_3}(P)}{\text{Taille}(C, o)}\right)^2}$$

$$\text{avec } C \in \{M, V\} \text{ et } P \in C$$

$$\text{CouleurMoyenne_Montagnes}(o) = CMM(o) = \frac{\sum_{i=1}^{NM(o)} \text{Couleur}(M_i, o)}{NM(o)}$$

$$\text{CouleurMoyenne_Vallées}(o) = CMV(o) = \frac{\sum_{j=1}^{NV(o)} \text{Couleur}(V_j, o)}{NV(o)}$$

$$\sigma_{TMM}(o) = \sqrt{\frac{\sum_{i=1}^{NM(o)} (Taille(M_i, o) - TMM(o))^2}{NM(o)}} \quad \sigma_{TMV}(o) = \sqrt{\frac{\sum_{j=1}^{NV(o)} (Taille(V_j, o) - TMV(o))^2}{NV(o)}}$$

$$\sigma_{CMM}(o) = \sqrt{\frac{\sum_{i=1}^{NM(o)} (Couleur(M_i, o) - CMM(o))^2}{NM(o)}} \quad \sigma_{CMV}(o) = \sqrt{\frac{\sum_{j=1}^{NV(o)} (Couleur(V_j, o) - CMV(o))^2}{NV(o)}}$$

$$Rapport(o) = \frac{NV(o) \cdot TMV(o) \cdot CMV(o)}{NM(o) \cdot TMM(o) \cdot CMM(o)}$$

Afin de mesurer la répartition des zones de Montagne et des zones de Vallée, on calcule la position des centres de gravité de chaque région M_i et V_j ainsi que le centre de gravité de l'objet o . On définit alors :

$$Répartition_Montagnes(o) = \frac{1}{NM(o)} \sum_{i=1}^{NM(o)} \frac{\sqrt{\pi} \cdot \sqrt{(\overline{x_{M_i}} - \overline{x})^2 + (\overline{y_{M_i}} - \overline{y})^2}}{Taille(M_i, o)}$$

$$Répartition_Vallées(o) = \frac{1}{NV(o)} \sum_{j=1}^{NV(o)} \frac{\sqrt{\pi} \cdot \sqrt{(\overline{x_{V_j}} - \overline{x})^2 + (\overline{y_{V_j}} - \overline{y})^2}}{Taille(V_j, o)}$$

où $\overline{x_C}$ désigne le centre de gravité d'une région C , et \overline{x} le centre de gravité de l'objet.

3.4.5. Chapeau haut de forme

Le chapeau haut de forme est une opération de morphologie mathématique qui consiste à utiliser la différence entre l'image érodée (notée E) et l'image originale. On consultera [Serra 82] pour plus de détails. Soit $I_{C_1C_2C_3}$ une image couleur contenant un objet o :

$$I_{TH} = \frac{(E(I_{C_1C_2C_3}^{C_1}) - I_{C_1C_2C_3}^{C_1}) + (E(I_{C_1C_2C_3}^{C_2}) - I_{C_1C_2C_3}^{C_2}) + (E(I_{C_1C_2C_3}^{C_3}) - I_{C_1C_2C_3}^{C_3}))}{3}$$

$$\sigma_{TH} = \frac{\sum (I_{TH}(p))^2 - \frac{(\sum I_{TH}(p))^2}{N}}{Surface(o)}$$

I_{TH} représente l'image de la luminance de la différence entre l'image couleur érodée et l'image couleur originale. I_{TH} est seuillée afin de diviser l'objet o en plusieurs zones décrivant la texture permettant d'obtenir les régions intéressantes. Le seuil est égal à l'écart type de I_{TH} .

Nous avons utilisé un élément structurant carré de taille 7×7 . Pour un objet provenant de la cytologie des séreuses, la Figure 48 présente l'image I_{TH} ainsi que l'image I_{TH} seuillée.

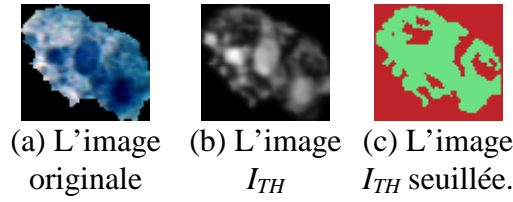


Figure 48. Les régions obtenues par la méthode du chapeau haut de forme dans l'espace RVB .

Notons N le nombre de régions r_i associées à l'objet o à caractériser, on définit pour chaque région :

$$Taille_TH(r_i, o) = Card\{P / I_s(P) = o \text{ et } P \in r_i\}$$

$$Couleur_TH(r_i, o) = \sqrt{\left(\frac{\sum I_{C_1 C_2 C_3}^{C_1}(P)}{Taille_TH(r_i, o)}\right)^2 + \left(\frac{\sum I_{C_1 C_2 C_3}^{C_2}(P)}{Taille_TH(r_i, o)}\right)^2 + \left(\frac{\sum I_{C_1 C_2 C_3}^{C_3}(P)}{Taille_TH(r_i, o)}\right)^2}$$

$$CouleurMoyenne_TH(o) = CM_TH(o) = \frac{1}{N} \sum_{i=1}^N Couleur_TH(r_i, o)$$

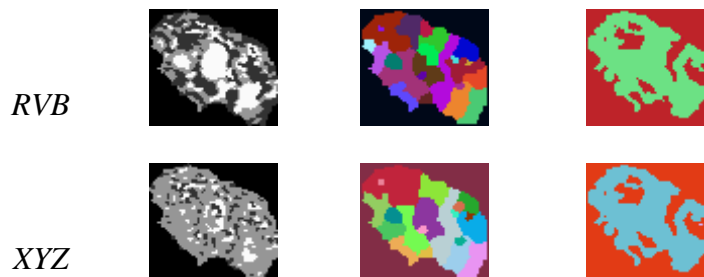
$$TailleMoyenne_TH(o) = TM_TH(o) = \frac{1}{N} \sum_{i=1}^N Taille_TH(r_i, o)$$

$$\sigma_{TM_TH}(o) = \sqrt{\frac{1}{N} \sum_{i=1}^N (Taille_TH(r_i) - TM_TH(o))^2}$$

$$\sigma_{CM_TH}(o) = \sqrt{\frac{1}{N} \sum_{i=1}^N (Couleur_TH(r_i) - CM_TH(o))^2}$$

3.4.6. L'espace couleur en analyse de texture

Pour toutes les méthodes d'extraction de texture que nous avons exposées dans les paragraphes précédents, nous n'avons illustré nos propos que par des images dans l'espace RVB . Une question se pose alors : un autre espace couleur est-il « utile » pour l'extraction de la texture ? Les trois méthodes (textons, L.P.E et chapeau haut de forme) ont été appliquées sur la cellule de la Figure 46.a et les résultats sont donnés par la Figure 49.



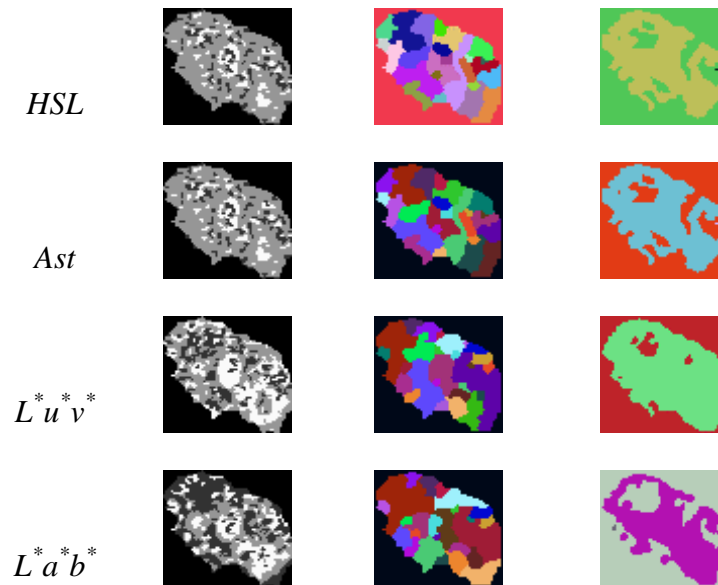


Figure 49. Les différentes méthodes d'extraction de texture dans plusieurs espaces couleur avec de gauche à droite les méthodes : textons, LPE et chapeau haut de forme.

On remarque immédiatement que pour la méthode d'extraction des textons, seul l'espace *RVB* semble être satisfaisant même si l'espace *L^{*}u^{*}v^{*}* paraît intéressant. Pour la méthode de ligne de partage des eaux et de chapeau haut de forme, les résultats se valent sauf pour les espaces *L^{*}u^{*}v^{*}* et *L^{*}a^{*}b^{*}* qui sont moins performants.

Il est très difficile d'analyser directement l'influence de l'espace couleur sur la caractérisation d'une texture. En effet seules les mesures faites pour représenter le concept de texture sont importantes et le seul moyen pour connaître l'apport d'un espace couleur dans la caractérisation d'un texture serait de classer toutes les caractéristiques afin de déterminer quel espace couleur est le mieux adapté.

Le problème de la pertinence d'une caractéristique est un problème très complexe que nous verrons plus en détail dans la partie suivante. Cependant nous avons décidé de ne pas chercher à utiliser nos caractéristiques de texture dans un autre espace. La raison vient principalement du fait que, comme pour les méthodes de segmentation, cela dépend du niveau de précision que l'on veut atteindre. Celui que l'on obtient dans l'espace *RVB* étant très satisfaisant, l'utilisation d'autres espaces couleur ne nous paraît pas indispensable.

4. Conclusion

Dans cette première partie de la thèse, nous nous sommes intéressés au problème de la segmentation d'images et de la caractérisation d'objets dans le domaine de la couleur. Nous avons tout d'abord abordé le problème de la segmentation. La ligne de partage des eaux étant particulièrement adaptée pour segmenter des images complexes telles que les images couleur, nous avons axé notre réflexion sur la définition d'un critère d'agrégation couleur. Ce critère prend en compte des informations locales (exprimées par le gradient couleur) et des informations globales (exprimées par l'homogénéité couleur d'une région). Afin de bien cerner les points importants de notre méthode de segmentation, nous avons défini une stratégie d'optimisation de la fonction d'agrégation. Plus précisément, nous avons proposé une méthode de choix de l'espace couleur basée sur le contraste couleur. L'utilisation et l'optimisation de notre méthode de segmentation couleur pour deux classes d'images (de cytologie et de nœuds de morceaux de bois) a permis de montrer son efficacité.

Une fois la segmentation d'une image réalisée, une étape importante d'un système de vision est la caractérisation des objets. Nous avons présenté un certain nombre d'attributs (dont certains sont classiques) en nous attachant principalement à l'étude de la texture. Les méthodes proposées de segmentation de la texture sont complémentaires et permettent d'avoir un certain point de vue multi-échelle de la texture en couleur. Chacune d'entre elles est appliquée directement sur l'image couleur : on ne fait pas un traitement composante par composante. Les méthodes de segmentation de texture sont utiles pour caractériser des objets mais la pertinence des caractéristiques reste à définir par une méthode de sélection d'attributs. Nous verrons dans la deuxième partie de la thèse comment sélectionner les attributs pertinents parmi ceux que nous avons sélectionnés.

5. Références

[Altunbasak 98] Altunbasak Y., Eren P.E., Tekalp A.M., « Region-Based Motion Segmentation Using Color Segmentation », *Graphical Models and Image Processing*, Vol. 60, n°1, 9. pp.13-23, 1998.

[Angot 98] Angot F., « *Segmentation d'images 2D et 3D : application à la quantification d'images histologiques et cytologiques obtenues par microscopie* », Thèse de l'Université de Caen, 1998.

[Belhomme 97] Belhomme P., Elmoataz A., Herlin P., Bloyet D., « Generalized region growing operator with optimal scanning : application to segmentation of breast cancer images », *Journal of Microscopy*, Vol. 186, pp. 41-50, 1997.

[Beucher 90] Beucher S., « *Segmentation d'images et morphologie mathématique* », thèse de l'école Nationale des Mines de Paris, 1990.

[Brock-Gunn 95] Brock-Gunn S.A., Dowling G.R., Ellis T.J., « Tracking Using Colour Information », *Internal Report*, 1995.

[Campadelli 97] Campadelli P., Medici D., Schettini R., « Color Image Segmentation using Hopfield Networks », *Image and Vision Computing*, Vol. 15, n° 3, pp.161, 1997.

[Carron 94] Carron T., Lambert P., « Color edge detector using jointly hue, saturation and intensity », *Proceedings of ICIP'94*, Vol. 3, pp. 977-981, 1994.

[Celenk 90] Celenk M., « A color clustering technique for image segmentation », *Computer Vision Graphics and Image Processing*, Vol. 52, pp.145-170, 1990.

[Chassery 84] Chassery J.M., « An iterative Segmentation Method based on contextual color and shape criterion », *IEEE transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, n°6, pp 794-800, 1984.

[CIE 31] Commission Internationale de l'Eclairage, *Proceedings of the International congress on Illumination*, Cambridge, Cambridge University Press,1931.

[CIE 76] CIE Publication n°15, Supplément numéro 2 (E-1.3.1), « Official recommendations on uniform Color Spaces, Color Difference Equations and Metric Color Terms », *Commission Internationale de l'Eclairage*, 1976.

[CIE 86] CIE Publication n°15.2, « Colorimetry, Second Edition (1986) », *Central Bureau of the Commission Internationale de l'Eclairage*, Vienne, Autriche, 1986.

[Clouard 99] Clouard R., « ATHENA », *Rapport Technique du GREYC*, Université de Caen, 1999.

[Comaniciu 97] Comaniciu D., Meer P., « Robust analysis of feature spaces : color image segmentation », *Proceedings of CVPR'97*, Puerto-Rico, 1997.

[Coster 89] Coster M., Chermant J.L., « *Précis d'analyse d'images* », Les Presses du CNRS, Paris, 1989.

[Daily 89] Daily M.J., « Color Image Segmentation Using Markov Random Fields », *Proceedings de Computer Vision and Pattern Recognition'89*, pp. 304-312, 1989.

[Delmotte 97] Delmotte D., « *Segmentation d'images à l'aide de réseaux de neurones* », Rapport de stage de D.E.A Intelligence Artificielle et Algorithmique, Université de Caen, 1997.

[Dibos 97] Dibos F., Koepfler G., « Segmentation d'images couleur par méthode variationnelle », *Proceedings du 16^{ème} colloque GRETSI*, pp. 367-370, 1997.

[Di Zenzo 86] Di Zenzo S., « A note on the gradient of a multi-image », *Computer Vision, Graphics and Image Processing*, Vol. 33, pp. 116-126, 1986.

[Elmoataz 96] Elmoataz A., Belhomme P., Herlin P., Schupp S., Revenu M., « Automated Segmentation of cytological and Histological Images for the nuclear Quantification : An adaptative Approach based on Mathematical Morphology », *Microsc. Microanal. Microstruct.*, Vol. 7, pp. 331-337, 1996.

[Faugeras 79] Faugeras O.D., « Digital Image Processing within the framework of a Human Visual Model », *IEEE transactions on Acoustic, speech and signal processing*, Vol. 27, n°7, pp. 380-393, 1979.

[Gaio 94] Gaio M., « *Une architecture d'aide à la construction d'interprétation géographique : contribution à l'interprétation des images aériennes* », Thèse de l'Université de Caen, 1994.

[Garbay 86] Garbay C., « Image Structure Representation and Processing : A Discussion of Some Segmentation Methods in Cytology », *IEEE transactions on Pattern Analysis and Machine Intelligence*, Vol. 8, n°2, pp. 140-146, 1986.

[Garcia-Bonafé 96] Garcia-Bonafé M., Moragas M., « Differential Diagnosis of malignant and reactive cells from serous effusions : image and texture analysis study », *Analytical and Quantitative Cytology and Histology*, Vol. 12, pp. 85-98, 1996.

[Gevers 90] Gevers T., Groen F.C.A., « Segmentation of color images », *Technical report*, University of Amsterdam, 1990.

[Giménez-Mas 94] Giménez-Mas J.A., Sans-Moncasi M.P., Remon L., Gambo P., Gallego-Calvo M.P., « Automated Textural Analysis of Nuclear Chromatin, A Mathematical Morphology Approach », *Analytical and Quantitative Cytology and Histology*, Vol. 17, n° 1, 1995.

[Guzmán de León 96] Guzmán de León A., Lerallut J-F., Boulanger J-C., « Segmentation of Colposcopic Color Images », *Medical & Biological Engineering & Computing*, Vol. 34, Supplement 1, Part 1, 1996.

- [Haralick 79] Haralick R.M., « Statistical and Structural Approaches to Texture », *Proceedings of the IEEE*, Vol. 67, n° 5, May 1979.
- [Healey 96] Healey C.G., Enns J.T., « A perceptual color Segmentation Algorithm », <<http://cs.ubc.ca/~healey>>, 1996.
- [Jackway 94] Jackway P.T., Deriche M., « Scale-space properties of the multiscale morphological dilatation-erosion », *IEEE transaction on Pattern Analysis and Machine Intelligence*, 1994.
- [Jackway 95] Jackway P.T., « Gradient watersheds in morphological scale-space », *Proceedings de IEEE Workshop on Nonlinear Signal and Image Processing*, Neos Maramas, Halkidiki, Greece, pp. 22-25, 1995.
- [Jackway 95b] Jackway P.T., « Scaled gradient watersheds and cell feature extraction », *DICTA'95 Digital Image Computing : Techniques and Applications*, pp. 234-239, Brisbane, Australia, 6-8 December 1995.
- [Julesz 81] Julesz B., « Textons, the elements of texture perception, and their interactions », *Nature*, Vol. 290, n° 12, pp. 91-97, 1981.
- [Jorgensen 96] Jorgensen T., Yosegan K., Tvetter K.J., Skjorten F, Danielsen H.E, « Nuclear Texture Analysis : A new Prognosis Tool in Metastatic Prostate Cancer », *Cytometry*, Vol. 24, pp. 277-283, 1996.
- [Kauppinen 95] Kauppinen H., Silven O., « A Color Vision Approach for Grading Lumber », *Theory & Applications of Image Processing II - Selected papers from the 9th Scandinavian Conference on Image Analysis*, ed. G. Borgefors, World Scientific, pp. 367-379, 1995.
- [Kunt 91] Kunt M., « *Traitement de l'information numérique des images* », Presses polytechniques et Universitaires romanes, 1991.
- [Lasserre 97] Lasserre P., Murrieta R., Briot M., « Le modèle Nominatif de Régions : segmentation couleur et identification de régions par analyse de couleur et de texture », *Proceedings du 16^{ème} colloque GRETSI*, pp. 39-42, 1997.
- [Levkovitz 93] Levkovitz H., Herman G.T., « GLHS : a Generalized Lightness, Hue and Saturation Color Model », *Graphical Models and Image Processing*, Vol. 55, n°4, pp. 271-285, 1993.
- [Littman 97] Littman E., Ritter E., « Colour Image Segmentation : a Comparison of Neural and Statistical Methods », *IEEE Transaction on Neural Networks*, Vol. 8, n°1, pp. 175-185, 1997.
- [Liu 1994] Liu J., Yang Y-H., « Multiresolution Color Image Segmentation », *IEEE transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, n°7, pp. 689-700, 1994.
- [Luong 90] Luong Q-T. « *La couleur en vision par ordinateur : 1-Une revue* », Rapport de recherche INRIA n°1251, 1990.

- [MacAulay 89] MacAulay C., Tezcan H., Palcic B., « Adaptive Color Basis Transformation, An Aid in Image Segmentation », *Analytical and Quantitative Cytology and Histology*, Vol. 11, n°1, pp. 53-58, 1989.
- [Meyer 92] Meyer F., « Color Image Segmentation », *Proceedings of the 4th Conference Image Processing and its Applications*, pp. 303-306, 1992.
- [Moragas 93] Moragas A., Garcia-Bonafé M., De Torres I., Sans M., « Textural Analysis of Lymphoid Cells in Serous Effusions, A Mathematical Morphologic Approach », *Analytical and Quantitative Cytology and Histology*, Vol. 15, n°3, 1993.
- [Nicole 97] Nicole D., <http://www.erag.fr/MAC/Base.htm>, Ecole Romane des Arts Graphiques, Lausanne, Suisse, Janvier 1997.
- [Ohta 80] Ohta Y-I., Kanade T., Sakai T., « Color Information for regions segmentation », *Computer Graphics and Image Processing*, Vol. 13, pp. 222-241, 1980.
- [Pappas 92] Pappas T.N., « An adaptative Clustering Algorithm for Image Segmentation » *IEEE Transactions on Signal Processing*, Vol. 40, pp. 901-914, 1992.
- [Pauli 76] Pauli H., « Proposed extension to the CIE recommendation on uniform color spaces, color difference equations and metric color terms », *J. Opt. Soc. Am.*, Vol. 66, n° 8, pp. 866-867, 1975.
- [Poynton 95] Poynton C.A., « Frequently Asked Questions about Color », <http://www.inforamp.net/~poynton>, 1995.
- [Pratt 78] Pratt W.K., « *Digital Image Processing* », John Wiley & Sons, New-York, 1978.
- [Raghu 95] Raghu P.P., Poongodi R., Yegnanarayana B., « A combined neural network approach for texture classification », *Neural Networks*, Vol. 8, pp. 975-987, 1995.
- [Robinson 98] Robinson G., « Segmentation via RGB clustering », <http://www.npac.syr.edu/copywrite/pcw/node431.html>, 1998.
- [Saarinen 94] Saarinen K., « Color Image segmentation by a watershed algorithm and region adjacency graph processing », *Proceedings of ICIP'94*, Vol. 3, pp. 1021-1025, 1994.
- [Saarinen 95] Saarinen K., « Color Image Segmentation », *Internal Report*, 1995.
- [Salembier 96] Salembier P., Brigger P., Casas J.R., Pardas M., « Morphological Operators for Image and Video Compression », *IEEE transactions on Image Processing*, Vol. 5, n° 6, pp. 881-897, 1996.
- [Sammouda 97] Sammouda R., Niki N., Nishitani H., « Hopfiels Neural-Network for the Multichannel Segmentation of Magnetic-Resonance Cerebral Images » *Pattern Recognition*, Vol. 30, n°6, pp. 921-927, 1997.

- [Schettini 93] Schettini R., « A segmentation algorithm for color images », *Pattern Recognition Letters*, Vol. 14, n°6, pp. 499, 1993.
- [Serra 82] Serra J., « *Image Analysis and Mathematical Morphology* », Academic Press, Londres, 1982.
- [Shafarenko 97] Shafarenko L., Petrou M., Kittler J., « Automatic Watershed Segmentation of Textured Color Images », *IEEE transactions on Image Processing*, Vol. 6, n°11, pp. 1530-1543, 1997.
- [Silven 96] Silven O., Kauppinen H., « Recent Developments in Wood Inspection », *International Journal on Pattern Recognition and Artificial Intelligence, IJPRAI*, pp. 83-95, 1996.
- [Smolle 96] Smolle J., « Optimization of Linear Combination for Segmentation in Red-Green-Blue Images », *Analytical and Quantitative Cytology and Hystology*, Vol. 18, n°4, pp. 323-329, 1996.
- [Soille 96] Soille P., « Morphological partitioning of multispectral images », *Journal of Electronical Imaging*, pp. 252-265, 1996.
- [Tomczak 98] Tomczak R., Auroyer C., Rouquet C., Bonton P., « Color image segmentation algorithm for automated pollens analysis », *Actes de TISVA '98*, Vol. 1, pp. 11-18, Oujda, Maroc, 1998.
- [Tremeau 93] Tremeau A., « *Contribution des modèles de la perception visuelle à l'analyse d'images couleur* », Thèse de l'Université de Saint-Etienne, 1993.
- [Tremeau 97] Tremeau A., Borel N., « A region growing and merging algorithm to color segmentation », *Pattern Recognition*, Vol. 30, n° 7, pp. 1191-1203, 1997.
- [Umbaugh 93] Umbaugh S.E., Moss R.H., Stoecker W.V., Hance G.A., « Automatic Color Segmentation Algorithms with Application to Skin Tumor Feature Identification », *IEEE Engineering in Medicine and Biology*, pp. 75-82, September 1993.
- [Vincent 90] Vincent L., Soille P., « Watersheds In Digital Spaces : An Efficient Algorithm Based on Immersion Simulations », *IEEE transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, n°6, pp. 583-598, 1991.
- [Vincent 93] Vincent L., « Morphological grayscale reconstruction in image analysis : Applications and efficient algorithms », *IEEE Transactions on Image Processing*, Vol. 2, n°2, pp. 176-201, 1993.
- [Walker 94] Walker R.F., Jackway P., Lovell B., Longstaff I.D., « Classification of cervical cell nuclei using morphological Segmentation and textural feature extraction », *Second Australian and New-Zealand Conference on Intelligent Information Systems*, 1994.
- [Walker 96] Walker R.F., Jackway P., « Statistical geometric features, extensions for cytological texture analysis », *ICPR'96*, 1996.

[Witkin 83] Witkin A.P., « Scale-space filtering », *Proceedings de International Joint Conference on Artificial Intelligence (IJCAI)*, Palo-Alto, pp. 1019-1022, 1983.

[Wolf 95] Wolf G., Beil M., Guski H., « Chromatin Structure Analysis Based on a Hierarchic Texture Model », *Analytical and Quantitative Cytology and Histology*, Vol. 17, n° 1, pp. 25-34, 1995.

[Wright 89] Wright W.A., « Markov Random Field Approach to data fusion and colour segmentation », *Image Vision Comput.*, Vol .7, pp. 144-150, 1989.

[Wyszecki 82] Wysecki G., Stiles W.S., « *Color Science : Concepts and Methods, Quantitative Data and Formulae, Second Edition* », John Wiley & Sons, New-York, 1982.

[Zugaj 98] Zugaj D., Lattuati V., « A new approach of color images segmentation based on fusing region and edge segmentation outputs », *Pattern Recognition*, Vol. 31, n°2, pp. 105-113, 1998.

Classification par réseaux de neurones

1. MONNA : UNE ARCHITECTURE DE RÉSEAUX DE NEURONES.....	103
1.1. INTRODUCTION AUX RÉSEAUX DE NEURONES.....	103
1.1.1. LE MODÈLE DE NEURONE FORMEL.....	103
1.1.2. LES RÉSEAUX MULTICOUCHES.....	104
1.1.3. PROBLÈMES LIÉS AUX RÉSEAUX MULTICOUCHES.....	107
1.1.3.1. Structure du réseau.....	107
1.1.3.2. Prétraitement des entrées	108
1.2. PROBLÉMATIQUE	108
1.3. PRINCIPE DE L'ARCHITECTURE.....	109
1.4. CONSTRUCTION D'UNE ARCHITECTURE : MONNA.....	109
1.5. CRÉATION DE L'ARCHITECTURE	110
1.6. APPRENTISSAGE DES RÉSEAUX DE NEURONES.....	110
1.6.1. PRÉTRAITEMENT DES DONNÉES	111
1.6.2. STRUCTURE ADAPTATIVE	111
1.6.3. DÉTAILS SUR L'APPRENTISSAGE.....	113
1.7. HIÉRARCHISATION DES RÉSEAUX DE NEURONES	115
1.7.1. NOTATIONS.....	116
1.7.2. LA SÉLECTION PAR ÉLIMINATION	117
1.8. EXPÉRIMENTATIONS DE L'ARCHITECTURE MONNA.....	118
1.8.1. DESCRIPTION DES BASES DE DONNÉES	118
1.8.1.1. Base Liver Bupa.....	118
1.8.1.2. Base Pima.....	118
1.8.1.3. Base Ionosphere	118
1.8.1.4. Base Cancer	118
1.8.1.5. Base Vehicle	118
1.8.1.6. Base Pageblocks.....	119
1.8.1.7. Base Glass.....	119
1.8.1.8. Base Segmentation.....	119
1.8.1.9. Base OptDigits	119
1.8.1.10. Base Serous	119
1.8.1.11. Récapitulatif.....	119
1.8.2. COMPARAISON GLOBALE ENTRE UN RMC ET MONNA.....	120
1.8.3. LA STRUCTURE ADAPTATIVE	121
1.8.4. LE NOMBRE DE NEURONES DE LA COUCHE CACHÉE	123
1.8.5. LA SÉLECTION PAR ÉLIMINATION	125
1.8.6. COMPARAISON GÉNÉRALE.....	127
2. SÉLECTION D'ATTRIBUTS	129
2.1. LE PROBLÈME DE LA SÉLECTION D'ATTRIBUTS	129
2.2. DÉFINITION DE LA PERTINENCE	130
2.3. COMPOSANTES DE BASE DES MÉTHODES DE SÉLECTION DE CARACTÉRISTIQUES.....	132
2.3.1. LES ESPACES DE RECHERCHE	135
2.3.2. LES FONCTIONS D'ÉVALUATION	136
2.4. APPROCHES FILTER ET WRAPPER	137
2.5. ALGORITHMES DE RECHERCHE.....	138
2.5.1. L'ALGORITHME <i>HILL-CLIMBING</i>	138
2.5.2. L'ALGORITHME <i>BEST-FIRST SEARCH</i>	139
2.5.3. L'ALGORITHME <i>SEQUENTIAL FORWARD FLOATING SELECTION (SFFS)</i>	139

2.6.	APPROCHES TESTÉES	140
2.7.	L'APPROCHE FILTER	140
2.7.1.	TRACE DES MATRICES DE COVARIANCE INTER ET INTRA CLASSE	140
2.7.2.	THÉORIE DE L'INFORMATION	141
2.7.3.	MÉTHODE DE YACOUB	142
2.8.	L'APPROCHE WRAPPER.....	142
2.9.	RÉSULTATS EXPÉRIMENTAUX	142
2.9.1.	COMPARAISON SUR LES TAUX DE RECONNAISSANCE.....	143
2.9.2.	COMPARAISON DES APPROCHES <i>FILTER</i>	144
2.9.3.	COMPARAISON DES APPROCHES <i>WRAPPER</i> SUR LE NOMBRE D'ATTRIBUTS	147
 3. CONCLUSION.....		148
 4. RÉFÉRENCES.....		151

Figures

FIGURE 1. LE MODÈLE DU NEURONE FORMEL (POUR UN NEURONE i).	103
FIGURE 2. DEUX FONCTIONS SIGMOÏDES.	104
FIGURE 3. LA STRUCTURE D'UN RMC.	104
FIGURE 4. LE OU EXCLUSIF : IMPOSSIBILITÉ DE SÉPARER LINÉAIREMENT.	104
FIGURE 5. LE PROBLÈME DES MINIMA LOCAUX.	107
FIGURE 6. SCHÉMA DE LA STRUCTURE DES RÉSEAUX DE NEURONES UTILISÉS DANS L'ARCHITECTURE MONNA.	110
FIGURE 7. UN EXEMPLE DE RÉSEAU DE NEURONES À STRUCTURE ADAPTATIVE QUAND UN NEURONE EST AJOUTÉ À LA COUCHE CACHÉE (EN GRIS), LES LIENS SONT MIS À JOUR ET LES POIDS SONT TOUS INITIALISÉS DE FAÇON ALÉATOIRE.	112
FIGURE 8. LES ENSEMBLES SERVANT À L'APPRENTISSAGE ET À LA MESURE DU TAUX DE RECONNAISSANCE D'UN RÉSEAU.	114
FIGURE 9. L'ARCHITECTURE MONNA DANS SON INTÉGRALITÉ.	116
FIGURE 10. VARIATION DU TAUX DE RECONNAISSANCE (SUR L'ENSEMBLE D'ÉVALUATION) EN FONCTION DU NOMBRE DE NEURONES DE LA COUCHE CACHÉE DES RÉSEAUX LIÉS À LA BASE "GLASS".	124
FIGURE 11. MOYENNE DE LA VARIATION DU TAUX DE RECONNAISSANCE (SUR L'ENSEMBLE D'ÉVALUATION) EN FONCTION DU NOMBRE DE NEURONES DE LA COUCHE CACHÉE DES RÉSEAUX LIÉS À LA BASE "GLASS".	125
FIGURE 12. LES DIFFÉRENTES COMPOSANTES D'UN PROCÉDÉ DE SÉLECTION D'ATTRIBUTS.	134
FIGURE 13. CHAQUE ÉTAT DE L'ESPACE DES ATTRIBUTS SPÉCIFIE LES ATTRIBUTS À UTILISER. CHAQUE ENFANT (À DROITE) POSSÈDE UN ATTRIBUT (RONDS NOIR) DE MOINS QUE SES PARENTS [BLUM 97].	136
FIGURE 14. LES DEUX APPROCHES UTILISÉES EN SÉLECTION D'ATTRIBUTS, LA DIFFÉRENCE SE SITUE AU NIVEAU DE L'INCORPORATION DU CLASSIFIEUR.	137
FIGURE 15. HISTOGRAMME COMPARANT LES DIFFÉRENTES APPROCHES DE SÉLECTION D'ATTRIBUTS.	144
FIGURE 16. ÉCARTS ENTRE LES DIFFÉRENTES MÉTHODES UTILISANT L'APPROCHE <i>FILTER</i> POUR LE TAUX DE RECONNAISSANCE GLOBAL DE L'ARCHITECTURE (A) ET LE NOMBRE D'ATTRIBUTS MOYEN EMPLOYÉS PAR CHAQUE RÉSEAU (B).	146
FIGURE 17. ÉCARTS ENTRE LES DIFFÉRENTES MÉTHODES UTILISANT L'APPROCHE <i>WRAPPER</i> POUR LE TAUX DE RECONNAISSANCE GLOBAL DE L'ARCHITECTURE (A) ET LE NOMBRE D'ATTRIBUTS MOYEN EMPLOYÉ PAR CHAQUE RÉSEAU (B).	148
FIGURE 18. COMPARAISON DES TAUX DE RECONNAISSANCE DE L'ARCHITECTURE MONNA AVEC ET SANS SÉLECTION D'ATTRIBUTS AINSI QU'AVEC UN RMC.	149

Tableaux

TABLEAU 1. LES BASES DE DONNÉES UTILISÉES POUR LES TESTS.....	120
TABLEAU 2. COMPARAISON ENTRE LE TAUX DE RECONNAISSANCE D'UN RMC ET DE MONNA SUR LES BASES DU TABLEAU 1.....	121
TABLEAU 3. LE GAIN EN TERME DE TAUX DE RECONNAISSANCE ENTRE UN RMC ET MONNA ET LE NOMBRE DE CLASSES.....	121
TABLEAU 4. COMPARAISON ENTRE LE NOMBRE DE NEURONES DE LA COUCHE CACHÉE ENTRE UN RMC ET MONNA.....	122
TABLEAU 5. QUALITÉ ET NOMBRE DE NEURONES DE LA COUCHE CACHÉE DES RÉSEAUX DE L'ARCHITECTURE MONNA SUR LA BASE "VEHICLE".....	123
TABLEAU 6. COMPARAISON DES TAUX DE RECONNAISSANCE POUR LES DEUX CRITÈRES DE SÉLECTION PAR ÉLIMINATION.....	126
TABLEAU 7. LE NOMBRE MOYEN DE RÉSEAUX DE NEURONES ASSOCIÉS À LA DÉCISION DE CLASSIFICATION SELON LA SÉLECTION UTILISÉE.....	127
TABLEAU 8. COMPARAISON ENTRE DIFFÉRENTES APPROCHES NEURONALES.....	128
TABLEAU 9. LA PERTINENCE DES ATTRIBUTS POUR LE PROBLÈME XOR CORRÉLÉ SELON LES QUATRES DÉFINITIONS DE LA PERTINENCE D'UN ATTRIBUT.....	131
TABLEAU 10. COMPARAISON DES DIFFÉRENTES MÉTHODES DE SÉLECTION D'ATTRIBUTS EMPLOYÉES.....	143
TABLEAU 11. NOMBRE DE CLASSES ET D'ATTRIBUTS POUR CHAQUE CLASSE SERVANT AUX EXPÉRIMENTATIONS.....	145
TABLEAU 12. TAUX DE RECONNAISSANCE ET NOMBRE D'ATTRIBUTS SÉLECTIONNÉS POUR LA MÉTHODE YACOB SUR LES BASES.....	145
TABLEAU 13. TAUX DE RECONNAISSANCE ET NOMBRE D'ATTRIBUTS SÉLECTIONNÉS POUR LA MÉTHODE COV SUR LES BASES.....	145
TABLEAU 14. TAUX DE RECONNAISSANCE ET NOMBRE D'ATTRIBUTS SÉLECTIONNÉS POUR LA MÉTHODE QUINLAN SUR LES BASES.....	146
TABLEAU 15. TAUX DE RECONNAISSANCE ET NOMBRE D'ATTRIBUTS SÉLECTIONNÉS POUR LA MÉTHODE SFS SUR LES BASES.....	147
TABLEAU 16. TAUX DE RECONNAISSANCE ET NOMBRE D'ATTRIBUTS SÉLECTIONNÉS POUR LA MÉTHODE SBS SUR LES BASES.....	147
TABLEAU 17. TAUX DE RECONNAISSANCE ET NOMBRE D'ATTRIBUTS SÉLECTIONNÉS POUR LA MÉTHODE SFFS SUR LES BASES.....	148
TABLEAU 18. COMPARAISON DES TAUX DE RECONNAISSANCE ENTRE L'APPROCHE MONNA ET D'AUTRES APPROCHES NEURONALES.....	149

Les réseaux de neurones sont souvent utilisés en classification de données. Dans cette partie, nous proposons une architecture de réseaux de neurones utile lorsque les données sont réparties suivant un grand nombre de classes. Après une introduction rapide sur les réseaux de neurones, nous détaillerons les fondements de l'architecture utilisée et nous verrons son application à la classification de données. Ensuite, nous aborderons le problème de la sélection d'attributs pertinents et nous comparerons différentes approches en les appliquant à notre architecture.

1. MONNA : une architecture de réseaux de neurones

1.1. Introduction aux réseaux de neurones

1.1.1. Le modèle de neurone formel

Nous allons décrire ici un modèle du neurone et par la suite nous décrivons un modèle de connexions entre ces neurones (les réseaux multicouches). Un neurone est une entité reliée à d'autres neurones et il peut être modélisé par deux opérateurs [Herrault 94] :

- Un opérateur de sommation qui élabore un état interne I_i égal à la somme pondérée des entrées de l'unité,
- Un opérateur d'activation qui calcule l'état de la sortie S_i du neurone en fonction de son état interne I_i .

D'une manière générale on peut définir un neurone formel par les cinq éléments suivants (Figure 1):

- La nature de ses entrées,
- La fonction de sommation du neurone qui définit son état,
- La fonction d'activation qui calcule la sortie du neurone,
- La nature de la sortie.

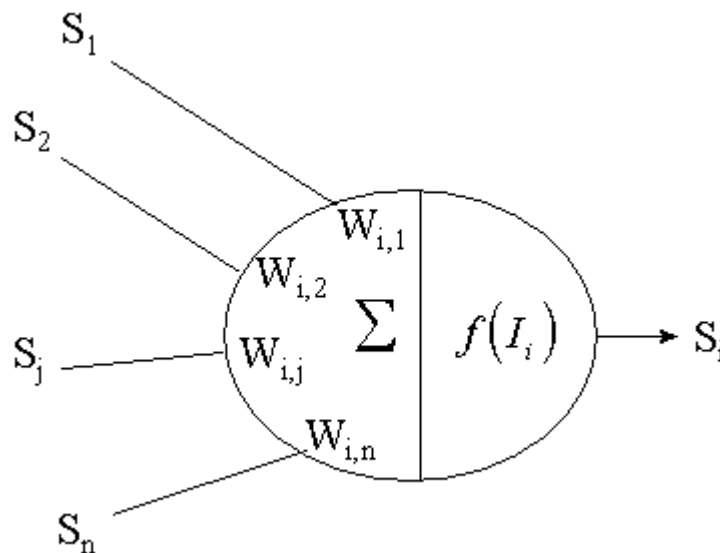


Figure 1. Le modèle du neurone formel (pour un neurone i).

Les calculs de l'état interne et de la sortie s'expriment par les relations suivantes :

$$I_i = \sum_{j=1}^n W_{i,j} \cdot S_j \text{ et } S_i = f(I_i) \text{ où } f \text{ est la fonction d'activation du neurone ou fonction de seuil.}$$

Dans le cas d'un réseau constitué de plusieurs neurones, ceux-ci sont repérés par un numéro

d'ordre i . $W_{i,j}$ désigne la valeur du poids de la connexion entre le neurone j et le neurone i , j désignant le neurone émetteur et i le neurone cible. Le processus d'activation d'un neurone se fait à travers une fonction de sommation (sommations pondérées des entrées), cette activation est ensuite transformée au travers d'une fonction de seuil. Les fonctions de seuil d'allure sigmoïde sont très utilisées : $f(x) = \frac{1}{1 + e^{-\beta x}}$. Ce sont des équations dont la forme générale est celle d'une tangente hyperbolique ou d'exponentielle avec des valeurs comprises entre -1 et $+1$ ou 0 et $+1$ (Figure 2).

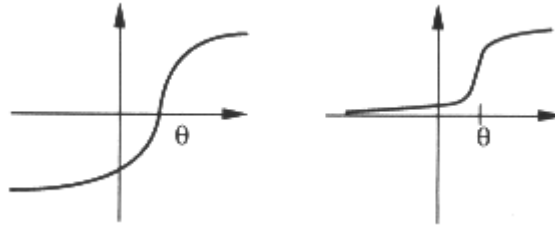


Figure 2. Deux fonctions sigmoïdes.

1.1.2. Les réseaux multicouches

Les réseaux de neurones auxquels nous nous intéressons sont des réseaux multicouches (RMC). Comme leur nom l'indique, ceux-ci sont organisés en couches et sont souvent appelés perceptrons multicouches dans le sens où ils sont une généralisation du perceptron qui ne possède qu'une couche [Minsky et Papert 69, Rumelhart 86]. Dans ce paragraphe, nous allons détailler la structure et l'apprentissage d'un RMC. On distingue trois sortes de couches composant un RMC : entrée, cachée ou sortie (Figure 3).

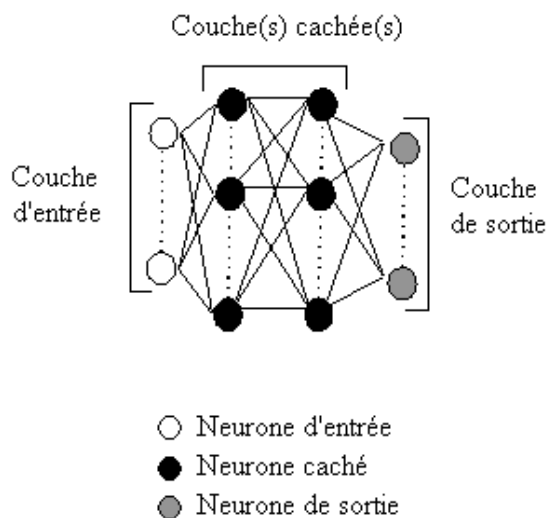


Figure 3. La structure d'un RMC.

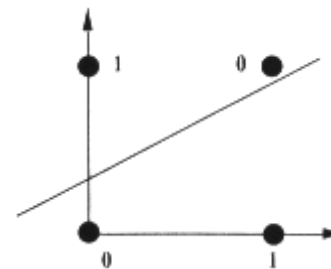


Figure 4. Le ou exclusif : impossibilité de séparer linéairement.

La couche d'entrée est la première couche, son rôle est de recevoir les vecteurs d'entrées, le nombre de neurones présents sur cette couche est donc égal à la dimension d'un vecteur d'entrée. En terme de classification de données, la couche d'entrée dispose d'autant de

neurones sur sa couche d'entrée que d'attributs décrivant un objet. Ces neurones agissent à la manière de tampons et n'ont d'autre rôle que de transmettre les entrées qui leur sont appliquées à la couche suivante. Ils constituent ce qu'il est d'usage d'appeler la rétine.

La ou les couches suivantes sont les couches cachées. En général, on insère de 0 à 2 couches cachées [Herrault 94], le terme de couche cachée désignant les neurones n'ayant aucune connexion avec l'extérieur du réseau. Un réseau possédant 0 couche cachée est limité à une séparation linéaire des données et est donc incapable de résoudre des problèmes tels que le ou exclusif. Cet exemple peut sembler sans intérêt mais il fût cependant à l'origine du manque d'intérêt suscité par les recherches neuronales dans les années 70. Le problème du ou exclusif suppose en effet une séparation non linéaire (Figure 4) que ne peut résoudre un réseau de type perceptron mis au point par Rosenblatt [Rosenblatt 60].

La couche de sortie est la dernière couche d'un RMC, elle contient autant de neurones que de sorties désirées. Habituellement on dispose d'autant de neurones de sortie qu'il y a de classes d'objets à discriminer.

Dans les réseaux multicouches, les connexions entre les couches de neurones sont totales et directes, c'est-à-dire que tous les neurones d'une couche sont reliés aux neurones de la couche suivante : chaque connexion transmet sa valeur aux différentes entrées de la couche supérieure. On ne trouve pas d'interconnexions à l'intérieur d'une même couche.

L'apprentissage d'un réseau multicouche se fait par l'algorithme de rétropropagation du gradient de l'erreur. Considérons un réseau à trois couches, la couche d'entrée ne fait que transmettre les entrées et le réseau ne possède que deux couches réelles de traitement : la couche cachée et la couche de sortie. Si l'on associe à un vecteur d'entrée un vecteur de sortie, étant donné les valeurs des poids des connexions, la sortie observée sera différente de celle souhaitée. A cette différence, on peut associer le terme d'erreur quadratique. L'apprentissage (supervisé) consiste à modifier les poids du réseau en minimisant cette erreur quadratique. Chaque présentation d'exemples différents provoquera la changement progressif des poids de connexion. Ces poids correspondent à la connaissance du réseau pour résoudre un problème particulier. Ils sont initialisés de façon aléatoire, puis vont s'adapter progressivement à partir d'exemples.

Définissons l'erreur quadratique sur les neurones de la couche de sortie. Pour un vecteur d'entrée x produisant des sorties S_i on connaît la sortie désirée Y_i :

$$E = \frac{1}{2} \sum_{i=1}^n (S_i - Y_i)^2, \text{ } i \text{ parcourant les neurones de sortie.}$$

Si l'on calcule le gradient de cette erreur par rapport au poids $W_{i,j}$ on obtient :

$$\frac{\partial E}{\partial W_{i,j}} = \sum_{k=1}^n (S_k - Y_k) \frac{\partial (S_k - Y_k)}{\partial W_{i,j}}$$

La dérivée partielle du membre droit est nulle sauf pour $k = i$, car seule la sortie S_i est fonction du poids $W_{i,j}$ ce qui donne :

$$\frac{\partial E}{\partial W_{i,j}} = (S_i - Y_i) \frac{\partial (S_i - Y_i)}{\partial W_{i,j}}$$

De plus la sortie désirée Y_i ne dépend d'aucun poids, on obtient finalement :

$$\frac{\partial E}{\partial W_{i,j}} = (S_i - Y_i) \frac{\partial S_i}{\partial W_{i,j}}$$

Or $S_i = f(I_i) = f\left(\sum_{j=1}^n W_{i,j} \cdot S_j\right)$ où f est la fonction de seuil, si bien que

$$\frac{\partial E}{\partial W_{i,j}} = (S_i - Y_i) \cdot f'(I_i) \cdot S_j \text{ où } f' \text{ est la dérivée de } f.$$

En posant $\delta_i = (S_i - Y_i) \cdot f'(I_i)$ le gradient de l'erreur quadratique E s'écrit :

$$\frac{\partial E}{\partial W_{i,j}} = \delta_i \cdot S_j.$$

Pour faire décroître l'erreur quadratique E , il faut modifier les poids de connexions proportionnellement à $\delta_i \cdot S_j$. La mise à jour de l'erreur quadratique selon le principe du gradient s'écrit :

$$\Delta W_{i,j} = -\alpha \cdot \delta_i \cdot S_j \quad (1)$$

où α est un gain d'adaptation positif en général assez petit (inférieur à 1).

Cependant, si cette méthode est directement applicable pour modifier les poids de la couche de sortie, elle ne peut pas s'appliquer directement aux couches internes. En effet, on ne connaît pas la sortie désirée pour ces couches, et par conséquent on ne connaît pas directement le terme d'erreur associé à chaque couche interne. Il faut donc tenter d'exprimer l'erreur à la sortie de chaque neurone d'une couche quelconque à partir de l'erreur de la dernière couche (seule erreur directement mesurable). Ceci est possible par un calcul différentiel et l'on peut montrer que l'erreur δ_i du neurone i est donnée par :

$$\delta_i = f'(I_i) \cdot \sum_{k=1}^n W_{k,i} \cdot \delta_k \text{ avec } k \text{ parcourant les neurones de la couche précédente.}$$

Ainsi la mise à jour des poids des connexions reliant les neurones entre eux peut se faire par une rétropropagation de l'erreur. Elle consiste à incriminer la totalité des neurones sans distinction en propageant l'erreur de sortie sur les différentes couches jusqu'à la couche d'entrée : le nom de rétropropagation du gradient de l'erreur tire son origine de cette méthode de distribution récurrente de l'erreur. La base de cet algorithme consiste à utiliser en remplacement d'une fonction de seuil linéaire, une fonction de seuil dérivable (fonction sigmoïde) qui mathématiquement se réduit à l'emploi des règles de dérivation composée. Pour résumer, la modification des poids lors de l'apprentissage supervisé se fait par la formule suivante :

$$W_{i,j} = W_{i,j} - \alpha \cdot \delta_i \cdot S_j$$

$$\text{avec } \begin{cases} \delta_i = (S_i - Sd_i) \cdot f'(I_i) & \text{pour la couche de sortie} \\ \delta_i = f'(I_i) \cdot \sum_{k=1}^n (W_{k,i} \cdot \delta_k) & \text{pour les autres couches} \end{cases}$$

L'algorithme de rétropropagation du gradient de l'erreur quadratique permet donc de faire évoluer les poids des connexions. Il présente cependant une importante limitation. On cherche à faire décroître l'erreur et pour ce faire on utilise la descente du gradient.

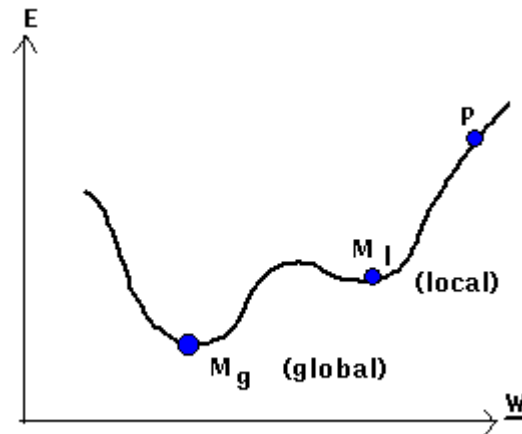


Figure 5. Le problème des minima locaux.

Supposons que l'on commence avec un ensemble de poids correspondant au point P . Si l'on utilise la descente du gradient, le minimum que l'on atteint est M_l et non pas M_g . M_l est un minimum local et correspond à une solution partielle du réseau de neurones en réponse aux données d'apprentissage. M_g est le minimum global que nous cherchons à atteindre et, à moins que certaines mesures soient prises pour échapper au minimum local M_l , M_g ne sera jamais atteint. Ce problème reste entier, on peut toujours utiliser des méthodes d'optimisation telles que le gradient conjugué mais rien ne nous assure que le minimum atteint est M_g , c'est-à-dire celui qui minimise la fonction d'erreur.

1.1.3. Problèmes liés aux réseaux multicouches

1.1.3.1. Structure du réseau

Pour des problèmes complexes où le nombre de classes est élevé et où les objets sont difficiles à différencier, l'obtention d'une structure optimale du réseau peut poser des problèmes. En effet, nous avons précisé qu'un réseau de neurones pouvait posséder plusieurs couches cachées et que celles-ci pouvaient contenir un nombre quelconque de neurones. Cependant nous n'avons pas encore abordé le problème du choix de la structure du réseau. Plusieurs approches théoriques ont montré que les réseaux de neurones à 3 couches (entrée, cachée et sortie) sont capables d'effectuer l'approximation de fonctions continues [Comon 91] de \mathbb{R}^p (un vecteur à p composantes) dans $[0,1]$. De façon plus simple et générale [Hérault 94], un réseau à 3 couches est capable d'approximer n'importe quelle fonction continue. On ne sait par contre pas comment construire le réseau, ni combien de neurones sont nécessaires sur la couche cachée. Le problème est que ce nombre de neurones peut être infini car certains problèmes font appel à une infinité de neurones avec un réseau à trois couches. Pour résumer le choix du nombre de neurones de la couche cachée est délicat : un nombre de neurones trop petit induira une modélisation insuffisante et un nombre de neurones trop grand entraînera une surparamétrisation du modèle, qui nuira aux performances en généralisation. Ceci pose un autre problème : celui de l'apprentissage. Pour obtenir de bons résultats, il faut trouver le bon nombre de neurones sur la couche cachée (souvent de façon tout à fait *ad hoc*). Or lorsque le nombre de neurones devient élevé, le nombre de connexions augmente lui aussi, et la difficulté de l'apprentissage du réseau s'en trouve augmenté. Or les réseaux de neurones sont généralement utilisés (et ne prennent de l'intérêt) que lorsque le problème est complexe, notamment lorsqu'il est non linéaire et mal connu. Si bien que tout l'intérêt d'utiliser des

réseaux s'en trouve amoindri par la détermination de sa structure qui pose un problème crucial : comment la déterminer.

1.1.3.2. *Prétraitement des entrées*

Si les neurones d'entrée reçoivent des valeurs de natures très différentes, il est évident que certaines de ces entrées n'auront pas la même influence dans l'apprentissage des poids du réseau. En pratique, dans un problème de classification, ceci pose un problème supplémentaire : la dynamique des entrées doit être de nature similaire afin de faciliter l'apprentissage (les entrées doivent appartenir au même type de données : réel ou binaire). En effet prenons deux variables x_1 et x_2 correspondant à deux entrées d'un réseau. Ces deux variables sont supposées de distribution uniforme :

$$-1 < x_1 < +1$$

$$-10 < x_2 < +10$$

Si l'on examine l'expression de la règle d'adaptation des poids du réseau (1), il est évident que les poids des liens reliant les neurones de la couche cachée à l'entrée x_2 varieront dix fois plus vite que ceux reliés à x_1 [Hérault 94]. Ceci prouve donc que les entrées doivent être prétraitées afin d'avoir des domaines de variation équivalents pour chacune des entrées. Ceci permet d'éviter l'influence de certaines entrées par rapport à d'autres.

1.2. *Problématique*

Nous nous plaçons dans le cadre de la classification de données par réseaux de neurones. Nous avons vu les fondements théoriques d'un RMC ainsi que le principe d'apprentissage. Les problèmes que nous avons évoqués concernant l'utilisation d'un RMC se situent principalement au niveau du potentiel de généralisation de celui-ci. En théorie, un RMC est capable d'approximer n'importe quelle fonction à la condition de disposer de suffisamment de neurones sur ses couches cachées, ce qui fait des RMC un classifieur universel. Lors de l'apprentissage, un RMC cherche à obtenir un taux de reconnaissance le plus élevé possible pour les objets de l'ensemble d'apprentissage. Ceci est a priori possible, cependant nous avons vu qu'imposer une structure très complexe à un RMC entraîne une surparamétrisation du modèle qui nuit aux performances de généralisation et donc à l'obtention d'un taux de reconnaissance optimal sur un ensemble de test. On distingue ici deux notions, celle d'apprentissage et celle de généralisation. Un réseau apprend bien lorsqu'il reconnaît de façon optimale les données de l'ensemble d'apprentissage et un réseau généralise bien lorsqu'il reconnaît de façon optimale des données de l'ensemble de test. Lorsque le problème à résoudre est complexe (au sens où les données sont difficiles à séparer), les réseaux prennent toute leur dimension, mais ils peuvent présenter des difficultés de généralisation. Ceci peut s'expliquer facilement. Pour un problème donné certaines classes peuvent être faciles à différencier (et peuvent parfois l'être de façon linéaire), mais d'autres classes peuvent être très difficiles à séparer et le réseau devra apprendre de façon suffisamment pertinente pour les séparer. Or le réseau doit reconnaître à la fois les classes facilement séparables et les classes difficilement séparables, ce qui entraîne un surcoût de l'apprentissage puisque le réseau devra posséder une structure complexe pour séparer correctement toutes les données. Le problème principal de l'utilisation des réseaux de neurones multicouches réside ici. Si l'on cherche à obtenir un apprentissage optimal, on risque de le faire au détriment de la généralisation et de plus cela entraîne une structure complexe du réseau.

Ce que nous allons chercher à réaliser, c'est simplifier la tâche du réseau de neurones afin que sa structure soit simple et son apprentissage moins complexe. Une simplification de la tâche du classifieur peut consister à simplifier l'espace des données à classer. Pour cela, il faut donc que le réseau ait moins de classes à discriminer. Or pour un nombre de classes n fixé, simplifier la tâche du RMC consiste à lui donner un nombre k de classes à discriminer tel que $k < n$. Ceci suppose donc d'utiliser plusieurs réseaux de neurones, chacun discriminant un nombre de classes faible par rapport au nombre total de classes. C'est cette hypothèse que nous proposons et dont nous allons montrer l'utilité et l'efficacité.

1.3. Principe de l'architecture

L'obtention d'un taux de reconnaissance élevé n'est pas une tâche triviale lorsque le problème à résoudre est complexe : ceci entraîne la détermination d'une structure appropriée pour un seul grand RMC. De plus, il n'existe pas de méthode efficace et générique déterminant la structure d'un réseau. Trouver une structure efficace peut requérir un grand nombre d'essais et de retours sur erreur qui doivent être menés par un spécialiste. L'approche proposée redéfinit le problème de la tâche d'apprentissage d'un RMC afin d'obtenir une règle simple de construction qui donnera de bons résultats de généralisation et ceci sans l'assistance d'un expert. Pour cela nous utilisons une stratégie de type « diviser pour régner ». L'objectif est de fractionner un problème de classification complexe en plusieurs problèmes plus simples. Cette répartition d'un problème de classification présente deux intérêts majeurs. D'une part elle permet de simplifier la tâche de classification demandée à chaque réseau et donc de simplifier l'obtention d'une structure appropriée lors de l'apprentissage et d'autre part elle doit permettre d'augmenter le taux de reconnaissance des données à classer. Cette simplification présente également la propriété de diminuer le temps d'apprentissage, en effet à problème de classification plus simple, structure plus simple et donc convergence plus rapide. Ce dernier point présente un avantage non négligeable pour des problèmes complexes présentant un nombre élevé de classes.

1.4. Construction d'une architecture : MONNA

Pour simplifier la tâche de classification d'un RMC, nous avons vu qu'il est nécessaire de diminuer le nombre de classes à différencier. Ceci entraîne l'utilisation de plusieurs réseaux. Il reste à déterminer quelles classes chaque réseau va discriminer. Ce que l'on veut réaliser, c'est simplifier au maximum la tâche de classification d'un réseau, nous avons donc choisi de n'utiliser que des réseaux servant à séparer deux classes. A chaque réseau est alors associée une décision simple : déterminer à laquelle des deux classes appartient un objet à classer. Si certaines classes sont linéairement séparables deux à deux, la structure du réseau associé à la séparation de ces deux classes sera simple et donc son potentiel de généralisation sera accru et son apprentissage plus court. Pour les classes qui ne sont pas linéairement séparables, la tâche de classification est plus complexe, mais comme un réseau ne distingue que deux classes, son potentiel de généralisation sera également accru car il ne se préoccupe que de la discrimination de ces deux classes. Dans ce dernier cas, le réseau parviendra dans la majorité des cas à une structure simple.

Le principe de notre architecture MONNA (Multiple Ordinate Neural Network Architecture) est donc simple et elle s'articule alors autour de trois phases :

1. La création de l'architecture suivant le nombre de classes à séparer,
2. L'apprentissage de chacun des réseaux de neurones,

3. La construction d'une hiérarchie de réseaux de neurones amenant à la classification des objets.

1.5. Création de l'architecture

La création de l'architecture de fait de façon automatique, elle nécessite de connaître à l'avance le nombre n de classes à distinguer. Nous devons alors créer pour n classes, η réseaux de neurones constituant l'architecture. Puisque nos réseaux ne sont destinés qu'à séparer deux classes, cela nous conduit à avoir $\eta = \frac{n \cdot (n-1)}{2}$ réseaux de neurones distincts.

L'architecture est donc constituée d'un ensemble disjoint de réseaux, il n'y a aucune relation entre les réseaux, on construit une forêt de réseaux. Une première remarque s'impose alors : lorsque le nombre de classes devient élevé, le nombre de réseaux devient très élevé. Par exemple pour un problème de séparation de 10 classes, cela amène à la création de $\frac{10 \times 9}{2} = 45$ réseaux de neurones. Ce nombre peut paraître élevé à première vue, mais grâce à la spécialisation de chaque réseau, la structure de chacun d'entre eux sera plus simple qu'un seul grand réseau de neurones. Le point faible de notre méthode est largement compensé par un taux de reconnaissance élevé et des apprentissages moins longs.

De plus, afin de simplifier la structure des η réseaux créés, nous allons utiliser pour chacun d'entre eux le synopsis suivant : une couche d'entrée composée d'autant de neurones que d'attributs décrivant un objet à classer, une seule et unique couche cachée composée d'un nombre variable de neurones et un seul neurone de sortie. Un seul neurone suffit : sa valeur étant comprise dans l'intervalle $]-1,1[$, suivant le signe de la sortie de cet unique neurone, un objet est classé dans l'une des deux classes (Figure 6).

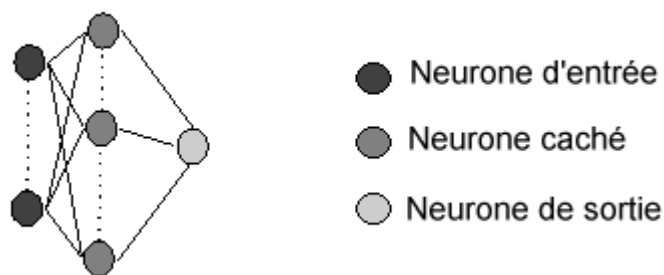


Figure 6. Schéma de la structure des réseaux de neurones utilisés dans l'architecture MONNA.

Lors de la création, on crée donc η réseaux en affectant de façon systématique deux classes différentes à différencier. On remarque que si le réseau séparant les classes C_1 et C_2 a été créé, on ne crée pas un réseau séparant les classes C_2 et C_1 (cela explique la division par deux dans le calcul de η).

1.6. Apprentissage des réseaux de neurones

La deuxième phase de création de l'architecture MONNA consiste à effectuer l'apprentissage de chacun des réseaux de neurones. L'apprentissage se fait de manière séquentielle et non ordonnée. Les réseaux apprennent un par un suivant l'ordre dans lequel ils ont été créés. Cela n'a aucune influence sur l'apprentissage de chacun des réseaux puisqu'il n'y a aucune connexion entre ceux-ci. Les réseaux sont conçus pour ne séparer que deux classes et par conséquent, il n'est utile de présenter à chaque réseau que des exemples de

données dont les classes correspondent à celles qu'il sépare. Comparativement à un seul grand RMC, cela présente un avantage puisque l'ensemble des données à apprendre s'en trouve restreint.

De plus, nous avons vu que nos réseaux possèdent une structure très simple (Figure 6), cependant nous nous trouvons confronté aux mêmes problèmes que pour la détermination de la structure d'un grand RMC : le nombre de neurones de la couche cachée nécessaires pour atteindre un taux de reconnaissance satisfaisant est inconnu. De même, il nous faut définir une stratégie de prétraitement des entrées.

1.6.1. Prétraitement des données

Afin d'éviter que les neurones d'entrée ne reçoivent des valeurs de natures très différentes, nous procédons avant l'apprentissage des réseaux à une normalisation des données d'entrée. Un objet est décrit par un certain nombre d'attributs qui ont le plus souvent des domaines de variation très différents. Nous allons réduire, pour chacun des attributs, leur domaine de variation à l'intervalle $[0,1]$. Cela se fait simplement grâce à une transformation linéaire qui correspond à une égalisation d'histogramme. Soit A la valeur d'un attribut donné, A_{max} et A_{min} les valeurs maximales et minimales que peut prendre cet attribut sur toutes les données d'apprentissage, on définit :

$$\rho(A) = \left(\frac{A - A_{min}}{A_{max} - A_{min}} \right)$$

Remplacer la valeur A de l'attribut par $\rho(A)$ ne suffit pas car les valeurs ne sont pas suffisamment étalées sur l'intervalle $[0,1]$. Pour pallier ce problème, on construit l'histogramme cumulé de ρ , et il suffit de prendre la valeur correspondante de $\rho(A)$ de l'histogramme cumulé. L'effet de cette transformation améliore la répartition des données et permet de simplifier l'apprentissage des réseaux car les données sont normalisées et sont par conséquent plus homogènes.

1.6.2. Structure adaptative

Le problème de l'organisation de la couche cachée de nos réseaux reste entier. Comme nos réseaux sont très simples (une seule couche cachée, un seul neurone de sortie) la généralisation de leur structure peut se faire de façon dynamique. En effet, puisque l'apprentissage des réseaux est relativement facilité par le fait qu'ils ne séparent que deux classes, il serait intéressant de déterminer automatiquement et de façon dynamique le nombre de neurones de la couche cachée. Quelques travaux de recherche ont été menés afin de construire dynamiquement des réseaux de neurones. Ceux-ci incluent la création dynamique des noeuds [Ash 89], l'algorithme « cascade correlation » [Fahlman 89], l'algorithme « tiling » [Mezard 89], l'algorithme « self-organizing » [Tenorio 90] et l'algorithme « upstart » [Frean 90]. Ces algorithmes sont conçus pour déterminer le nombre de neurones de la couche cachée avant l'apprentissage du réseau et donc indépendamment de la phase d'apprentissage. Mais si le nombre de neurones (qui est donc déterminé à l'avance) ne s'avère pas approprié, cela peut pénaliser l'apprentissage.

Travaillant sur des réseaux très simples, un apprentissage dynamique s'effectuant pendant l'apprentissage semble être une meilleure solution. Nous proposons donc de modifier la structure d'un réseau de neurones lors de la phase d'apprentissage afin de déterminer automatiquement et de manière optimale, le nombre de neurones de la couche cachée. Le principe est très simple : le réseau apprend avec des structures différentes, chaque structure ne différant de la précédente que par son nombre de neurones de la couche cachée. Le réseau

commence avec la structure la plus simple possible (correspondant à une séparation linéaire) c'est-à-dire un seul neurone sur la couche cachée. Des neurones sont ensuite ajoutés progressivement un à un à la couche cachée. Le but est d'obtenir la structure la plus simple donnant le meilleur taux de reconnaissance sur l'ensemble d'apprentissage. Chaque fois qu'un nouveau neurone est ajouté, le réseau procède à un nouvel apprentissage par la méthode de rétropropagation du gradient. Les poids initiaux sont obtenus de façon aléatoire à chaque ajout d'un nouveau neurone à la couche cachée (Figure 7).

Précisons ici la nature des ensembles de données, nous avons utilisé la stratégie suivante : l'ensemble initial contenant toutes les données est partitionné en deux : 80% comme ensemble servant à l'apprentissage (E_I) et 20% comme ensemble de test (E_2). L'ensemble d'apprentissage est à nouveau partitionné en deux ensembles : 80% comme l'ensemble des données à apprendre (E_{Ia}) et 20% comme l'ensemble des données d'évaluation (E_{Iv}). On dispose finalement de trois ensembles : un ensemble d'apprentissage (E_{Ia}), un ensemble d'évaluation (E_{Iv}) et un ensemble de test (E_2).

Lors de l'apprentissage, les données apprises sont celles de l'ensemble E_{Ia} mais le taux de reconnaissance mesuré lors de l'apprentissage l'est sur l'ensemble d'évaluation E_{Iv} .

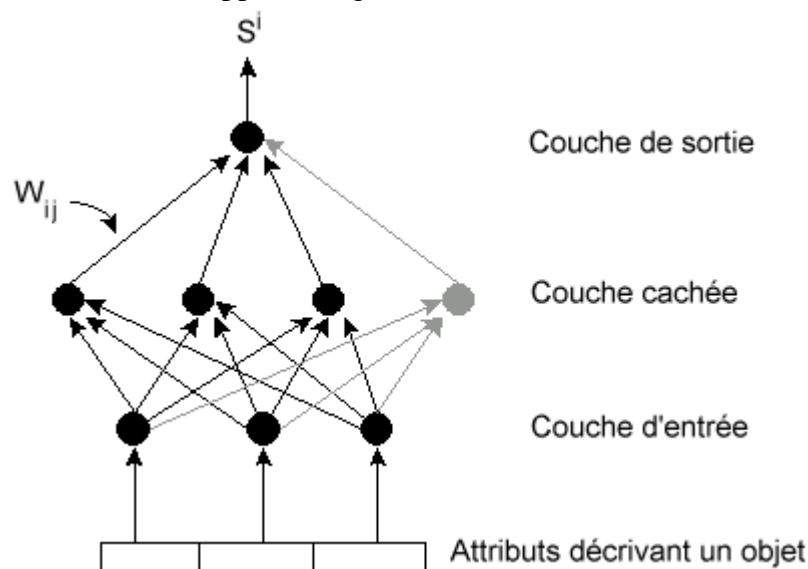


Figure 7. Un exemple de réseau de neurones à structure adaptative quand un neurone est ajouté à la couche cachée (en gris), les liens sont mis à jour et les poids sont tous initialisés de façon aléatoire.

L'apprentissage se termine soit lorsque le réseau a atteint un taux de classification satisfaisant sur les données de l'ensemble d'évaluation, soit lorsque le nombre maximal de neurones de la couche cachée a été atteint. On notera ici que pendant l'apprentissage on conserve toujours la structure du réseau possédant la structure la plus simple et qui a donné le meilleur taux de classification, car ce dernier peut augmenter puis décroître lors de l'ajout de neurones. L'algorithme d'apprentissage peut être décrit de la façon suivante pour un seul réseau de neurones.

Soient h le nombre de neurones de la couche cachée d'un réseau, h_{max} le nombre de neurones maximum de la couche cachée, Q_{max} le taux de reconnaissance satisfaisant, $R(h)$ le réseau ayant h neurones sur sa couche cachée, $Q(h)$ le taux de reconnaissance du réseau $R(h)$ sur les données de l'ensemble d'évaluation, R_{best} le réseau dont la structure amène au meilleur taux de classification et Q_{best} le taux de reconnaissance du réseau R_{best} . Le taux de reconnaissance est toujours mesuré sur l'ensemble d'évaluation.

1. $h = 1, R_{best} = R(1), Q_{best} = 0$
2. Initialiser les poids du réseau $R(h)$ de façon aléatoire.
3. Apprendre les poids par rétropropagation du gradient.
4. Si $Q(h) \geq Q_{max}$ alors fin de l'apprentissage et aller à l'étape 7.
5. Si $Q(h) > Q_{best}$ alors $Q_{best} = Q(h)$ et $R_{best} = R(h)$.
6. On ajoute un neurone à la couche cachée : $h = h + 1$. Si $h \leq h_{max}$ aller à l'étape 2.
7. La structure finale du réseau est celle de R_{best} .

L'algorithme de structure adaptative permet d'obtenir de façon fiable, le réseau de neurones qui apprend le mieux avec le moins de neurones cachés sur la couche cachée. Cette dernière condition est très importante car notre souci majeur concerne la simplification de la structure d'un réseau et le fait qu'un réseau possède trop de neurones cachés est indésirable. Si deux réseaux donnent le même taux de reconnaissance, il faut choisir celui qui est le moins complexe (moins de liens entre les neurones, donc moins de neurones sur la couche cachée). Il est toujours possible de construire un réseau avec suffisamment de neurones cachés pour que toutes les données de la base d'apprentissage soient correctement classées.

1.6.3. Détails sur l'apprentissage

Chacun des réseaux de neurones de l'architecture MONNA utilise l'algorithme de structure adaptative donné au paragraphe précédent. Cependant, nous n'avons pas précisé comment s'effectuait l'apprentissage pour une structure de donnée et donc avec un h fixé (point 3 de l'algorithme de structure adaptative). Tout d'abord la fonction d'erreur utilisée est l'erreur quadratique $E = \frac{1}{2} (S_i - Y_i)^2$ dans une version synchrone (les poids sont modifiés à chaque présentation d'un objet de l'ensemble d'apprentissage). $Y_i \in]-1,1[$ désigne la valeur désirée du neurone de sortie (la classe de l'objet à apprendre), S_i désigne la valeur du neurone de sortie, $S_i = f\left(\sum_j W_{ij} O_j\right)$ où W_{ij} désigne les poids entre le neurone de sortie (i) et les neurones de la couche cachée (j). O_j est la valeur de la sortie du neurone j de la couche cachée. La fonction de transfert est $f(y) = \frac{2}{(1 + e^{-\alpha y})} - 1$.

On notera que par la suite si un réseau de neurones sépare deux classes C_1 et C_2 alors un objet est considéré comme de la classe C_1 si $S_i < 0$ et C_2 si $S_i \geq 0$.

Une fois l'apprentissage d'un réseau pour un h donné effectué, on dispose de $Q(h)$, le taux de reconnaissance du réseau. Plutôt que d'évaluer $Q(h)$ directement sur les données apprises, ce terme est obtenu à partir d'un ensemble d'évaluation. Lorsqu'un réseau évalue son taux de reconnaissance, celui-ci est donc mesuré sur l'ensemble des données d'évaluation et ne concerne par conséquent que des données non apprises (Figure 8). De plus, lorsque le réseau apprend les données de l'ensemble d'apprentissage, celles-ci sont obtenues de façon aléatoire en alternant les deux classes à reconnaître. Ceci permet de ne pas privilégier une classe par rapport à une autre, surtout lorsque l'une des deux classes est représentée par moins d'éléments que l'autre.

L'apprentissage se termine soit lorsque l'on a atteint un nombre maximal d'apprentissages, soit lorsque l'on a atteint le taux de reconnaissance Q_{max} (qui est le même que pour l'algorithme de la structure adaptative).

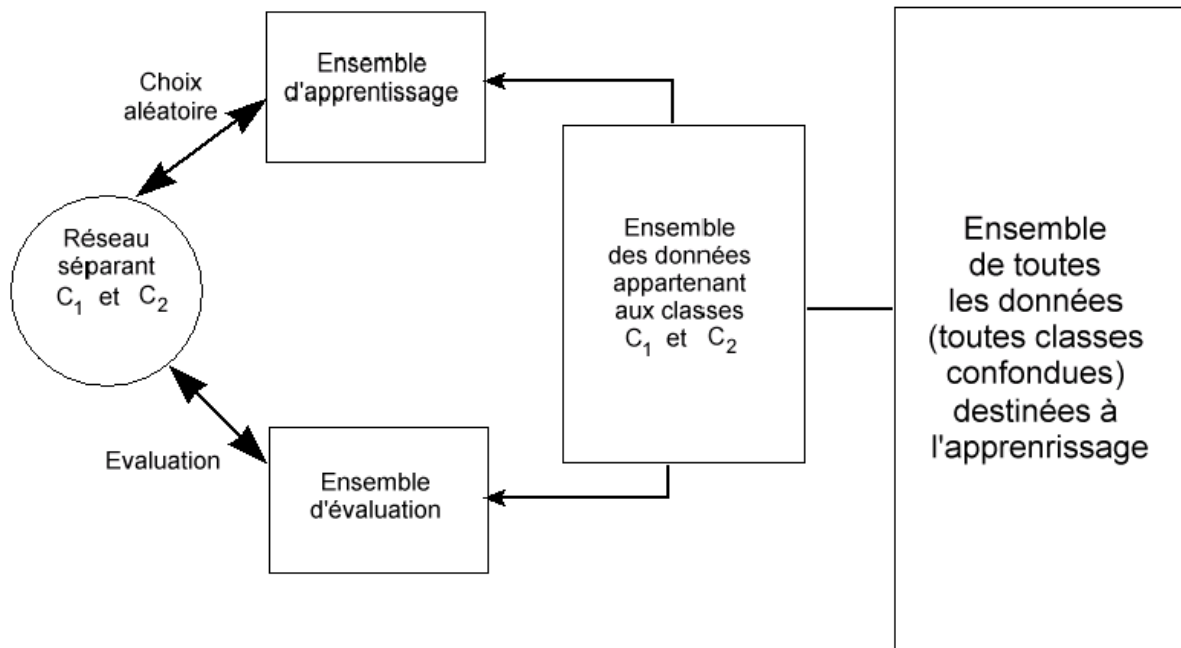


Figure 8. Les ensembles servant à l'apprentissage et à la mesure du taux de reconnaissance d'un réseau.

L'algorithme d'apprentissage pour un h donné (l'étape 3 de l'algorithme de structure adaptative) est:

```

NbreApp=0
Tant que NbreApp < NbreAppmax et  $Q(h) < Q_{max}$ 
  NbreApp = NbreApp + PasApp
  Pour  $n=0$  à PasApp
    si  $g$  est égal à  $C_1$ 
      alors  $g=C_2$ 
      sinon  $g=C_1$ 
    Tirage aléatoire d'un objet du groupe  $g$ 
      dans l'ensemble d'apprentissage
    Propager les valeurs
    Rétropropager les valeurs
    Modifier les poids
    Evaluer le réseau sur l'ensemble d'évaluation :
       $\Rightarrow$  modification de  $Q(h)$ 
  FinPour
  PasApp=1,2*PasApp
Fin tant que
  
```

1.7. Hiérarchisation des réseaux de neurones

La troisième phase de la création l'architecture MONNA consiste en la création d'une hiérarchie de réseaux de neurones permettant la classification des objets. Les deux premières phases ont permis de disposer d'un ensemble de η réseaux de neurones séparant chacun deux classes ainsi que de définir l'apprentissage des réseaux permettant de déterminer dynamiquement le nombre de neurones de la couche cachée. A ce stade lorsque l'on veut classer un objet, on peut disposer des réponses de chacun des réseaux. Cependant, il reste à déterminer comment utiliser ces différentes réponses pour déterminer la classe d'appartenance d'un objet. C'est le problème que nous allons évoquer lors de cette troisième phase. Nous allons créer une hiérarchie de réseaux de neurones permettant d'instaurer une relation d'ordre entre tous les réseaux. Cette hiérarchie se fait à plusieurs niveaux. Tout d'abord, le fait d'utiliser plusieurs réseaux de neurones séparant chacun uniquement deux classes permet d'avoir une vision hiérarchique des données, mais également une vision hiérarchique de la décision de classification qui s'affinera au fur et à mesure des interrogations successives des réseaux. La hiérarchie de réseaux de neurones repose sur trois points essentiels :

- On dispose de $\eta = \frac{n \cdot (n-1)}{2}$ réseaux de neurones séparant uniquement deux classes,
- Les réseaux sont ordonnés suivant un critère d'ordonnement,
- La décision de classification est prise par élimination.

Le premier point est le fondement même de notre architecture MONNA : chaque réseau n'est utilisé que pour séparer deux classes C_1 et C_2 . Quelle que soit la classe d'appartenance de l'objet à identifier, celui-ci sera classé dans l'une des deux classes. Si l'objet à classer n'appartient à aucune des deux classes, il sera classé sans aucune signification dans l'une ou l'autre classe.

Le deuxième point concerne la création de la hiérarchie de réseaux de neurones. Nous allons ordonner les réseaux suivant un critère qui va définir une hiérarchie. Deux critères ont été définis.

Le premier utilise le taux de reconnaissance des réseaux de neurones. Ce taux mesure la capacité d'un réseau à correctement classer un objet entre deux classes. On peut donc, à l'issue de l'apprentissage, classer les réseaux selon leur taux de reconnaissance (ou qualité de classification). Cette fiabilité est liée au fait que, le réseau ayant bien appris et réalisé une bonne généralisation, il est probable que, si les données des ensembles d'apprentissage et d'évaluation sont suffisamment représentatives, le réseau réalisera des classifications correctes.

Le deuxième critère défini pour former la hiérarchie utilise la valeur absolue du produit du taux de reconnaissance avec la valeur du neurone de sortie obtenue pour classer un objet : $|S(\mathcal{R}_i) \times Q(\mathcal{R}_i)|$. Pour le premier critère, on ne tient compte que du taux de reconnaissance obtenu lors de l'apprentissage. Cependant, la décision liée à la classification faite par un réseau dépend de deux facteurs : son taux de reconnaissance et la valeur du neurone de sortie (comprise entre -1 et $+1$). Si cette valeur est proche de 0 , cela correspond à une décision incertaine. Afin d'éviter ce cas de figure, qui mène à utiliser les réseaux ayant le mieux appris mais qui peuvent prendre des décisions incertaines, nous avons défini ce deuxième critère. Cela permet de choisir non pas le réseau qui a le mieux appris, mais celui qui a bien appris et qui classe de façon certaine (au sens de la valeur du neurone de sortie) un objet dans l'une des deux classes. La différence majeure entre les deux critères réside dans le fait que le deuxième utilise les sorties de tous les réseaux pour construire la hiérarchie. Nous

venons, grâce à cette étape, de réaliser une hiérarchie de réseaux de neurones qui va nous permettre de réaliser la classification des données.

Le troisième point définit plus exactement l'obtention de la classe d'un objet à classifier, c'est-à-dire la décision de classification. Chacun des réseaux permettant de décider à laquelle des deux classes un objet appartient, on dispose alors de η résultats qui peuvent être différents. Une manière de faire consisterait à choisir la classe de plus grande fréquence, mais cela pose un problème. En effet, si l'on propose à un réseau une donnée qui n'appartient à aucune des deux classes discriminées par celui-ci, la réponse n'est pas significative et cela risque de fausser le choix. Nous avons donc mis en place une sélection par élimination. Par sélection on entend détermination de la classe d'un objet et par élimination on désigne la façon de guider le choix de la classe finale. Si l'on décide d'utiliser un réseau de la hiérarchie, celui-ci va classer l'objet qui lui est présenté dans l'une des deux classes qu'il différencie (C_1 et C_2). Si le réseau désigne l'objet comme appartenant à C_1 alors C_2 est éliminée et réciproquement. C'est le principe de l'élimination. Cependant, il reste à définir comment réaliser le choix et l'enchaînement des réseaux pour réaliser la décision de classification. Pour cela, nous avons défini deux critères qui permettent chacun, mais de façon différente, de trouver la classe d'un objet à partir de différents réseaux de la hiérarchie.

L'architecture est ainsi totalement précisée. Sa construction, l'apprentissage de chacun des réseaux de neurones, ainsi que la décision de classification ont été définis. Nous pouvons maintenant schématiser l'architecture dans son intégralité (Figure 9).

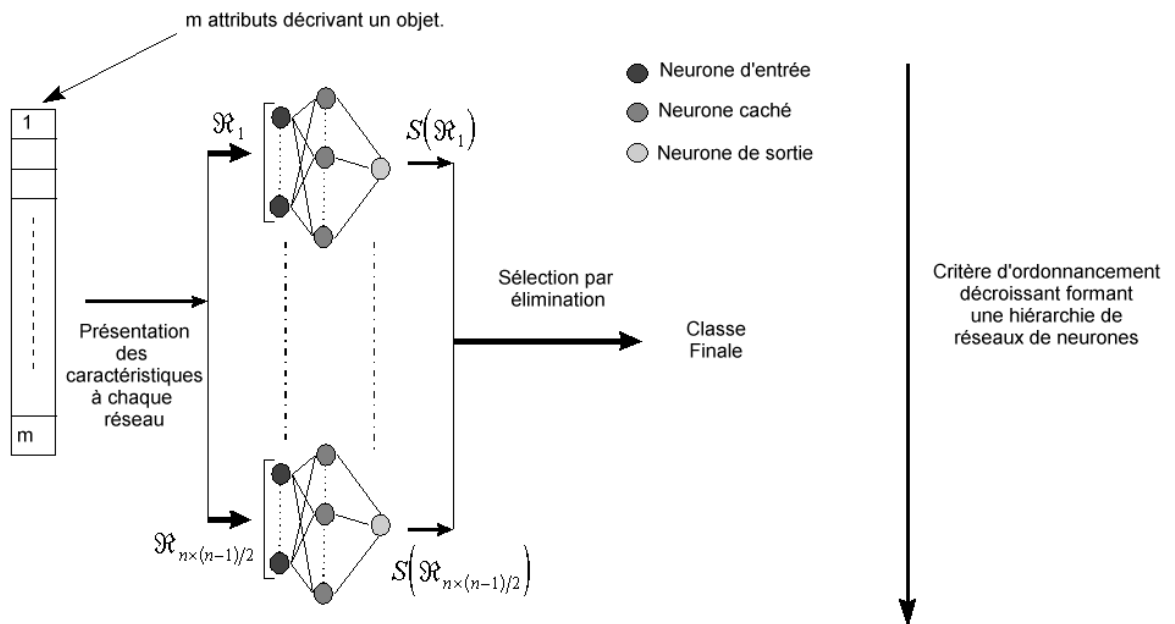


Figure 9. L'architecture MONNA dans son intégralité.

Notons que la façon dont est conçue l'architecture permet la prise en compte de nouvelles données en conservant une partie des réseaux. Cela peut permettre de modifier localement l'architecture : l'apprentissage de nouvelles données ne concerne que les réseaux qui sont concernés par ces nouvelles données. Ceci permet une nette amélioration par rapport à un seul grand RMC qui ne possède qu'un seul ensemble d'apprentissage.

1.7.1. Notations

Soit $\xi = \{C_1, C_2, \dots, C_n\}$ l'ensemble des classes des objets de la base de données, C_i désignant la classe d'un objet et n le nombre de classes. \mathfrak{R}_k désigne le réseau de neurones

numéro k dans la hiérarchie, \mathfrak{R}_k^p désigne la $p^{\text{ième}}$ classe discriminée par le réseau, on a $\mathfrak{R}_k^p \in \xi$ et $p \in \{1,2\}$ puisqu'un réseau ne discrimine que deux classes. Enfin, on définit $S(\mathfrak{R}_k)$ comme la valeur du neurone de sortie du réseau \mathfrak{R}_k , $\mathfrak{S} = \{\mathfrak{R}_1, \dots, \mathfrak{R}_{n \times (n-1)/2}\}$ comme l'ensemble des réseaux de neurones et $Q(\mathfrak{R}_k)$ comme le taux de reconnaissance du réseau numéro k .

1.7.2. La sélection par élimination

Quel que soit le critère d'ordonnement utilisé ($Q(\mathfrak{R}_k)$ ou bien $|S(\mathfrak{R}_k) \times Q(\mathfrak{R}_k)|$), le principe consiste à utiliser en premier le réseau dont le critère est le plus élevé. Celui-ci permet d'éliminer une classe. Ensuite, on passe au réseau suivant dans la hiérarchie et on continue jusqu'à ne plus avoir qu'une seule classe non éliminée. Cette classe est proposée par l'architecture MONNA pour l'objet présenté.

Sélection numéro 1:

$k = 1$

tant que $Card(\xi) \neq 1$

Si $S(\mathfrak{R}_k) < 0$

alors $\xi = \xi - \{\mathfrak{R}_k^1\}$

sinon $\xi = \xi - \{\mathfrak{R}_k^2\}$

$k = k + 1$

fin tant que

On peut également définir une autre sélection par élimination qui tient compte des classes déjà éliminées avant d'utiliser un réseau de la hiérarchie.

Sélection numéro 2:

$k = 1$

tant que $Card(\xi) \neq 1$

Si $S(\mathfrak{R}_k) < 0$

alors $Tmp = \mathfrak{R}_k^1$

sinon $Tmp = \mathfrak{R}_k^2$

Si $Tmp \in \xi$ alors $\xi = \xi - \{Tmp\}$

$k = k + 1$

fin tant que

1.8. Expérimentations de l'architecture MONNA

Afin de valider notre architecture de réseaux de neurones et de prouver son efficacité sur des données présentant un grand nombre de classes, nous avons procédé à différentes expérimentations. Les tests réalisés ici portent sur une large variété de bases de données dont les domaines d'origine sont parfois très différents (médicaux, segmentation d'images, reconnaissance de caractères, de vins, de voitures...). Nous avons utilisé 10 bases provenant de la collection de bases de l'université d'Irvine, Californie [Murphy 94] ainsi qu'une base supplémentaire résultant de nos propres travaux [Lezoray 98,99a,99b]. Dans un premier temps nous allons décrire chacune des bases de données que nous avons utilisé, puis nous procéderons à une étude détaillée des performances de l'architecture MONNA sur ces bases de données.

1.8.1. Description des bases de données

1.8.1.1. Base Liver Bupa

Cette base provient du centre de recherche médical BUPA (Richard S. Forsyth, Mapperley Park, Nottingham, UK). Les données sont réparties suivant deux classes et chaque vecteur de données est constitué de 5 attributs portant sur une personne de sexe masculin. Ces cinq variables correspondent à des tests sanguins donnant la sensibilité des troubles du foie qui peuvent apparaître lors d'une consommation excessive d'alcool.

1.8.1.2. Base Pima

Cette base provient du National Institute of Diabetes and Digestive and Kidney Diseases du Johns Hopkins University. Les données [Smith 88] sont réparties suivant deux classes et chaque vecteur de données est constitué de 8 attributs. Le problème consiste à prédire si un patient peut être positif à un test de dépistage du diabète selon les critères de l'O.M.S. (Organisation Mondiale de la Santé). Les données correspondent à des mesures pathologiques et à des résultats de tests médicaux.

1.8.1.3. Base Ionosphere

Cette base [Sigillito 89] provient du National Institute of Diabetes and Digestive and Kidney Diseases (USA). Les données sont réparties suivant deux classes et chaque vecteur de données est constitué de 34 attributs. Les données ont été collectées à partir d'un système de radars. Il est constitué d'un ensemble de 16 antennes de haute fréquence dont la puissance totale transmise est de l'ordre de 6.4 kilowatts. Les cibles sont des électrons dans l'ionosphère. Les signaux renvoyés sont soit bons soit mauvais. Les bons montrent la présence d'une structure dans l'ionosphère et les mauvais sont ceux qui passent à travers l'ionosphère.

1.8.1.4. Base Cancer

Cette base provient de l'hôpital Madison de l'Université du Wisconsin Les données sont réparties suivant deux classes et chaque vecteur est constitué de 9 attributs. Le problème consiste à différencier deux classes de cellules : anormale ou normale [Mangasarian 90].

1.8.1.5. Base Vehicle

Cette base provient du Turing Institute, Glasgow, Scotland. Les données sont réparties suivant 4 classes et chaque vecteur de données est constitué de 18 attributs. Le problème abordé ici consiste à distinguer des silhouettes 3D dans une image 2D en appliquant un ensemble d'opérateurs extracteurs de forme à la silhouette de chacun des objets 2D. Les

silhouettes utilisées sont des silhouettes de véhicules (un bus, un van Chevrolet, une Saab 900 et une Opel Manta 400). Cette combinaison particulière de véhicules a été choisie avec l'hypothèse que le bus et le van seraient facilement distingués des autres véhicules mais qu'il serait plus difficile de faire la distinction entre les deux voitures. Les véhicules ont subi plusieurs rotations et un certain nombre d'images (de silhouettes) ont été acquises en utilisant des angles de vue différents.

1.8.1.6. Base Pageblocks

Cette base de données [Malerba 95] provient de l'Université de Bari, Italie. Les données sont réparties suivant 5 classes et chaque vecteur de données est constitué de 10 attributs. Le problème consiste à classifier tous les blocs présents sur un document et détectés par un processus de segmentation. Ceci représente une étape essentielle dans l'analyse de documents afin de séparer le texte des zones de graphique. Les cinq classes correspondent à des zones de : texte, ligne horizontale, image, ligne verticale, graphique.

1.8.1.7. Base Glass

Cette base de données provient du Central Research Establishment, Aldermaston, Berkshire (USA). Les données sont réparties suivant 7 classes et chaque vecteur de données est constitué de 9 attributs. Le problème consiste à classifier des morceaux de verre en vue d'investigations de criminologie. Les morceaux de verre laissés lors de crime peuvent alors être utilisés comme preuves s'il sont bien identifiés.

1.8.1.8. Base Segmentation

Cette base provient du Vision Group, University of Massachusetts. Les données sont réparties suivant 7 classes et chaque vecteur de données est constitué de 19 attributs. Le problème consiste à classifier chacun des pixels d'images couleur de zones d'extérieur. Les images ont été segmentées manuellement afin d'affecter chaque pixel à l'une des classes suivantes : brique, ciel, feuillage, ciment, fenêtre, route et herbe.

1.8.1.9. Base OptDigits

Cette base de données [Kaynak 1995] provient de l'Université de Bogazici, Istanbul, Turquie. Les données sont réparties suivant 9 classes et chaque vecteur de données est constitué de 64 attributs. Le problème consiste à classifier des images normalisées de chiffres manuscrits. Les images noir et blanc (32x32 pixels) sont divisées en blocs de 4x4 pixels ne se chevauchant pas. Le nombre de pixels noirs est compté pour chaque bloc et il génère une matrice 8x8 où chaque élément est un entier appartenant à l'intervalle [0-16].

1.8.1.10. Base Serous

Cette base de données [Lezoray 98, 99a, 99b] provient des résultats de l'application des méthodes de la partie « Segmentation et caractérisation couleur » de cette thèse. Les données sont réparties suivant 18 classes et chaque vecteur de données est constitué de 46 attributs. Le problème consiste à classifier des cellules provenant de la cytologie des séreuses (cavités du corps humain telles que la plèvre et le péritoine). Ces cellules ont été segmentées et caractérisées à partir d'images couleur. Les attributs extraits portent sur des critères de texture, de couleur, de surface et de forme.

1.8.1.11. Récapitulatif

Nous disposons donc d'un ensemble de bases de données de provenances très diverses. Comme nous souhaitons tester notre architecture pour des problèmes présentant un grand

nombre de classes, les bases Wine, Vehicle, PageBlocks, Glass, Segmentation, OptDigits et Segmentation sont très intéressantes. Cependant, nous avons sélectionné des bases ne comportant que deux classes afin de comparer notre algorithme de structure adaptative avec un RMC classique. Le tableau 1 résume les propriétés des différentes bases que nous venons de citer.

Base	Nombre de classes	Nombre d'attributs	Nombre d'instances
Liver Bupa	2	6	345
Pima	2	8	768
Ionosphere	2	34	350
Cancer	2	9	683
Wine	3	13	178
Vehicle	4	18	846
PageBlocks	5	10	5473
Glass	7	9	214
Segmentation	7	19	210
OptDigits	9	64	3823
Serous	18	46	4837

Tableau 1. Les bases de données utilisées pour les tests.

Nous allons étudier les performances de l'architecture MONNA sur les bases que nous avons présentées. Pour chacune des bases 20% des données sont utilisées comme ensemble de test et 80% comme ensemble d'apprentissage. Ce dernier ensemble étant découpé à nouveau lors de l'apprentissage (voir § 3.5.3). Ce paragraphe s'articule de la façon suivante : tout d'abord nous allons comparer les performances globales de reconnaissance entre un RMC et l'architecture MONNA. Puis nous étudierons l'apport de l'algorithme de structure adaptative ainsi que le nombre de neurones utilisés sur la couche cachée. Nous comparerons pour finir les deux critères de sélection par élimination. Les conditions expérimentales sont les suivantes : Q_{max} est fixé à 95% et h_{max} est fixé à 50. Ces valeurs sont arbitraires et elles peuvent évidemment être modifiées, cependant nous avons constaté, par l'expérience qu'elles permettaient d'obtenir de bons résultats.

1.8.2. Comparaison globale entre un RMC et MONNA

Nous avons défini l'architecture MONNA afin d'améliorer les performances d'un classifieur neuronal sur différentes bases de données. Les performances d'un RMC et de MONNA se comparent directement grâce à leurs taux de reconnaissance respectifs. Nous avons procédé à l'apprentissage d'un RMC et de MONNA sur toutes les bases du § 1.8.1. Quelques bases de données (LiverBupa, Pima, Ionosphere et Cancer) présentent un nombre de classes à différencier égal à 2. Nous nous trouvons alors dans un cas particulier de notre architecture. En effet, pour $n = 2$, on a $\eta = 1$, ce qui revient à un RMC. Donc pour les quatre premières bases de données, les taux de reconnaissance d'un RMC et de MONNA seront identiques. On notera que nous avons utilisé l'algorithme de structure adaptative pour l'apprentissage d'un RMC (son nombre de neurones de la couche cachée sera donc déterminé de façon automatique).

Base	RMC	MONNA	Nombre de Classes
Liver Bupa	71%	71%	2
Pima	76,6%	76,6%	2
Ionosphere	90,1%	90,1%	2
Cancer	97,1%	97,1%	2
Wine	97,3%	100%	3
Vehicle	71,9%	78,4%	4
PageBlocks	85,4%	90,1%	5
Glass	66,7%	82,2%	7
Segmentation	82,4%	87,8%	7
OptDigits	78,1%	81,6%	9
Serous	55,9%	65,8%	18

Tableau 2. Comparaison entre le taux de reconnaissance d'un RMC et de MONNA sur les bases du tableau 1.

Le tableau 2 donne le taux de reconnaissance d'un RMC et de l'architecture MONNA sur les ensembles de test obtenus à partir des bases de données. On constate que dès que le nombre de classes devient supérieur à deux, MONNA permet d'obtenir de meilleurs résultats. En particulier pour la base de données Wine pour laquelle la classification obtenue est parfaite. Ceci confirme donc que notre architecture s'avère utile pour la classification de données lorsque le nombre de classes à différencier devient élevé. Cependant la différence constatée dans le taux de reconnaissance entre un RMC et MONNA ne semble pas être fonction du nombre de classes et dépend en grande partie de la base de données puisque que le gain lié au taux de reconnaissance ne dépend pas de façon linéaire du nombre de classes (Tableau 3).

Base	Gain	Nombre de Classes
Wine	2,7%	3
Vehicle	6,5%	4
PageBlocks	4,7%	5
Glass	15,5%	7
Segmentation	5,4%	7
OptDigits	3,5%	9
Serous	9,9%	18

Tableau 3. Le gain en terme de taux de reconnaissance entre un RMC et MONNA et le nombre de classes.

Nous pouvons constater tout de même que l'architecture MONNA permet d'obtenir un gain de reconnaissance compris entre 3 et 15% sur ces 7 différentes bases (Tableau 3), ce qui n'est pas du tout négligeable. Donc, lorsque le nombre de classes devient élevé, il semble préférable, afin d'obtenir un meilleur taux de reconnaissance d'utiliser notre architecture.

1.8.3. La structure adaptative

Nous venons de constater que l'architecture MONNA s'avère utile pour obtenir un bon taux de reconnaissance des données lorsque le nombre de classes devient élevé. Un autre avantage de l'architecture MONNA réside dans la simplicité des réseaux de neurones utilisés.

Cela provient directement de la structure de chacun des réseaux de neurones qui déterminent, lors de l'apprentissage, le nombre optimal de neurones de la couche cachée. Ceci fait la grande force de notre architecture : l'utilisation de réseaux possédant des structures simples.

Si l'on compare les structures d'un RMC et des réseaux de l'architecture MONNA en terme de nombre de neurones sur la couche cachée, on peut s'apercevoir que globalement les réseaux utilisés par MONNA sont plus simples. Le tableau 4 résume ces résultats. Pour le RMC est présenté le nombre de neurones de la couche cachée et pour MONNA le nombre minimal et maximal, la moyenne et l'écart-type du nombre de neurones de la couche cachée et ceci pour tous les réseaux utilisés. La structure des réseaux utilisés par MONNA est plus simple, mais ceci n'est vrai que lorsque que chaque réseau peut généraliser suffisamment facilement. Si ce n'est pas le cas, cela peut entraîner la création de réseaux complexes. Pour la base de données « *Serous* », on a un réseau qui a 46 neurones sur sa couche cachée, soit 11 de plus qu'un RMC. Cependant, même si ce réseau utilisé par MONNA présente plus de neurones sur la couche cachée relativement à un RMC, sa structure n'est pas plus complexe. Le nombre de liens dans le réseau utilisé par MONNA est de $46 \times 46 + 46 \times 1 = 2162$ et de $46 \times 35 + 46 \times 18 = 2240$ pour un RMC.

Base	RMC	MONNA			
		Min	Max	Moyenne	Ecart-Type
Liver Bupa	3	3	3	3	0
Pima	23	23	23	23	0
Ionosphere	7	7	7	7	0
Cancer	1	1	1	1	0
Wine	2	1	1	1	0
Vehicle	47	1	28	6,6	9,9
PageBlocks	31	1	17	2,8	4,7
Glass	25	1	11	1,8	2,3
Segmentation	18	1	1	1	0
OptDigits	47	1	1	1	0
Serous	35	1	46	4,5	10,3

Tableau 4. Comparaison entre le nombre de neurones de la couche cachée entre un RMC et MONNA.

Sur la base « *Vehicle* », on remarque (tableau 4) qu'un réseau de neurones possède 28 neurones sur sa couche cachée, ce qui peut paraître élevé. Ce réseau est dédié à la séparation des classes 0 et 1 (tableau 5), or ces classes sont difficiles à séparer si bien que cela entraîne un réseau plus complexe. En effet, si les classes sont difficiles à séparer, cela induit un taux de reconnaissance faible (inférieur à 60% ici) et donc un apprentissage plus long car la qualité Q_{max} n'est jamais atteinte et l'algorithme déterminant la structure optimale du réseau va continuer jusqu'au nombre de neurones maximal avant de s'arrêter. En général, on peut dire que certains réseaux nécessitent un apprentissage plus long, cela entraîne le parcours de tout l'ensemble des réseaux possible ($h=1$ à h_{max}). Dans certains cas de petits réseaux ont été conservés et dans d'autres cas de plus grands. Cet apprentissage n'utilise, pour un réseau de neurones, que les objets relatifs aux classes qu'il sépare, ce qui permet d'accélérer la convergence étant donné la simplicité de la tâche attribuée au réseau.

Neurones cachés	Qualité	Réseau Classe i, Classe j
1	96,4%	Réseau 2<->3
1	96,4%	Réseau 1<->3
5	96,5%	Réseau 0<->2
4	94,3%	Réseau 1<->2
1	93,9%	Réseau 0<->3
28	58,6%	Réseau 0<->1

Tableau 5. Qualité et nombre de neurones de la couche cachée des réseaux de l'architecture MONNA sur la base "Vehicle".

L'algorithme de structure adaptative permet donc d'obtenir des réseaux plus simples. On constate que pour toutes les bases, la majorité des réseaux de neurones ne possèdent qu'un nombre très réduit de neurones sur la couche cachée et ce dernier est souvent égal à un. Nous allons maintenant nous attarder de façon plus précise sur le nombre de neurones de la couche cachée.

1.8.4. Le nombre de neurones de la couche cachée

Nous avons vu au paragraphe précédent que l'algorithme de structure adaptative permet de déterminer automatiquement le nombre de neurones de la couche cachée et que ceci permet de simplifier la structure d'une grande partie des réseaux de neurones utilisés. Nous allons maintenant étudier l'influence de ce nombre de neurones sur l'apprentissage. En effet si l'algorithme de structure adaptative permet de minimiser le nombre de neurones, nous n'avons pas étudié la variation, en terme de taux de reconnaissance, qu'engendre cette recherche exhaustive.

L'étude de cette recherche exhaustive peut se faire en étudiant la variation du taux de reconnaissance d'un réseau sur l'ensemble d'évaluation d'une base de données. Nous avons choisi la base « Glass » pour effectuer nos expérimentations et ceci pour plusieurs raisons. Premièrement le nombre de classes à différencier est élevé (7) mais cela n'engendre pas un nombre de réseaux de neurones trop important (21), ce qui nous permet d'évaluer assez facilement les résultats. Deuxièmement le taux de reconnaissance global d'un RMC est peu élevé comparativement à celui de MONNA, ce qui préfigure un problème de classification assez difficile.

Pour un réseau \mathcal{R}_k séparant deux classes \mathcal{R}_k^p avec $p \in \{1,2\}$, on stocke $Q(h)$ lors de l'apprentissage. $Q(h)$ désigne le taux de reconnaissance du réseau avec h neurones sur sa couche cachée. On dispose ainsi d'une vision globale de l'évolution de l'apprentissage lorsque le nombre de neurones de la couche cachée augmente. La figure 10 présente cette variation du taux de reconnaissance en fonction du nombre de neurones de la couche cachée.

On constate que la majorité des réseaux atteignent un taux de reconnaissance de 100% avec seulement un neurone sur la couche cachée : seuls 5 réseaux ont une structure plus complexe qu'un seul neurone sur la couche cachée. Pour les 16 autres réseaux, puisque leur couche cachée ne présente qu'un seul neurone, le problème est linéairement séparable. Sur la figure 10, ne sont présentés que les réseaux pour lesquels la recherche de la meilleure configuration de la couche cachée se fait au delà d'un seul neurone.

On remarque pour chacun des réseaux de la figure 10, une grande variation du taux de reconnaissance lorsque le nombre de neurones de la couche cachée augmente. Cette variation paraît anormale. En effet, en théorie un RMC est capable d'approximer n'importe quelle fonction à la condition de disposer de suffisamment de neurones sur ses couches cachées, si bien que le taux de reconnaissance devrait augmenter lorsque le nombre de neurones sur la couche cachée augmente. Ce n'est pas le cas, comme le prouve la figure 10. Ces variations sont liées à deux facteurs : l'initialisation aléatoire des réseaux et l'obtention de minima locaux.

Les poids des réseaux de neurones sont initialement obtenus de façon aléatoire puis ces poids sont modifiés de façon à minimiser l'erreur quadratique. Cependant, l'initialisation aléatoire est différente à chaque nouvelle configuration du réseau car l'ajout d'un nouveau neurone provoque l'initialisation aléatoire de tous les poids. Dans certains cas, l'initialisation peut entraîner l'obtention d'un minimum local de l'erreur quadratique et le réseau de neurones termine son apprentissage prématurément. Afin de vérifier ces hypothèses quand à la variation du taux de reconnaissance, nous avons décidé de calculer une moyenne des taux de reconnaissance. Pour l'apprentissage d'un réseau avec un h fixé, on procède à 15 apprentissage successifs. Le taux de reconnaissance attribué au réseau pour un h donné est la moyenne des 15 taux de reconnaissance obtenus. La figure 11 présente la variation du taux de reconnaissance en fonction du nombre de neurones de la couche cachée, en utilisant cette dernière technique.

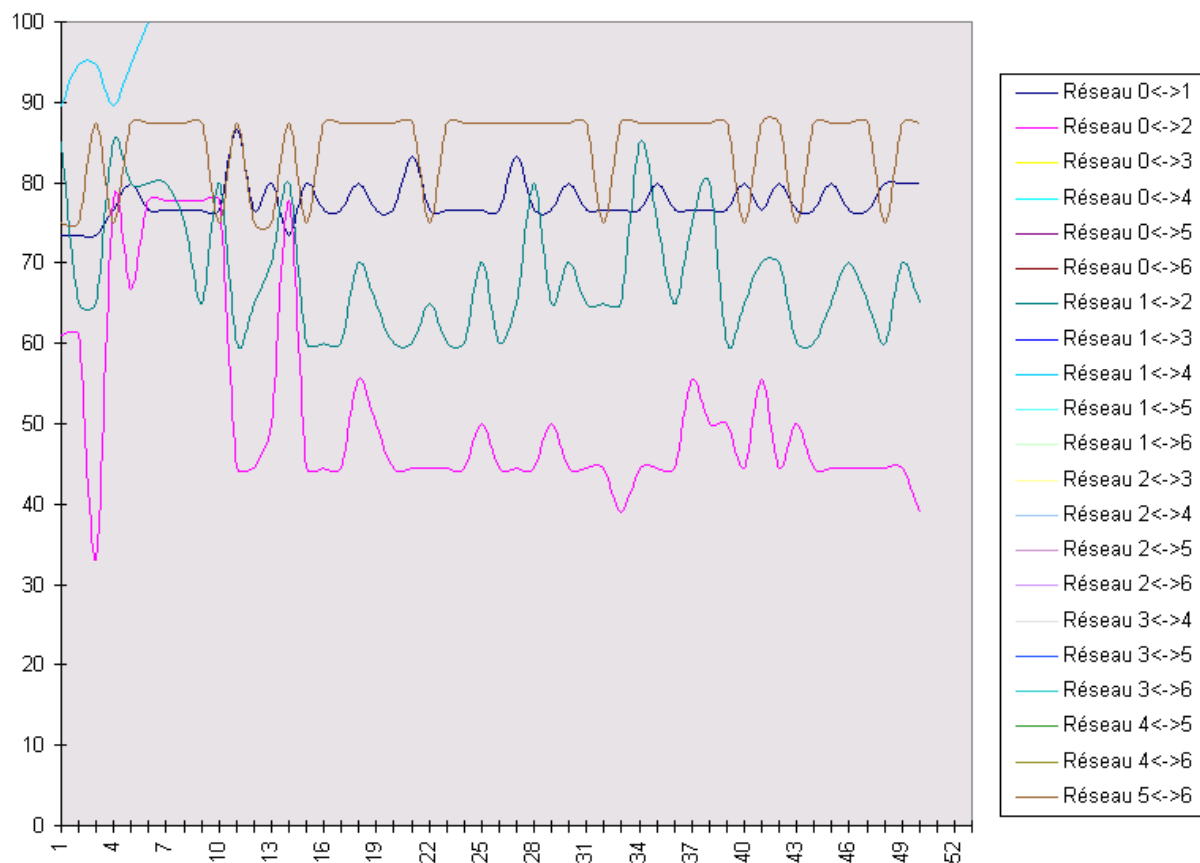


Figure 10. Variation du taux de reconnaissance (sur l'ensemble d'évaluation) en fonction du nombre de neurones de la couche cachée des réseaux liés à la base "glass".

On constate qu'effectivement moins de variations sont présentes et que bien que certains minima locaux subsistent, l'amplitude des variations est beaucoup moins élevée.

L'utilisation d'une moyenne de différents apprentissages pour une structure donnée permet donc d'avoir une meilleure vision de l'évolution du taux d'apprentissage en fonction du nombre de neurones de la couche cachée.

Sur les courbes des figures 10 et 11, les taux de reconnaissance optimaux sont atteints pour un nombre de neurones peu élevé de neurones sur la couche cachée. Lors de l'apprentissage effectif des différents réseaux de l'architecture MONNA, nous n'utilisons pas la moyenne de plusieurs apprentissage. Car si l'on compare les courbes des figures 10 et 11, on constate que même si plus de variations de grande amplitude sont présentes lorsque l'on n'utilise pas la moyenne, le taux de reconnaissance optimal est le même. De plus la mise en place de la moyenne de plusieurs apprentissage entraîne un surcoût trop élevé en temps d'exécution, si bien que cela n'a pas été retenu pour MONNA.

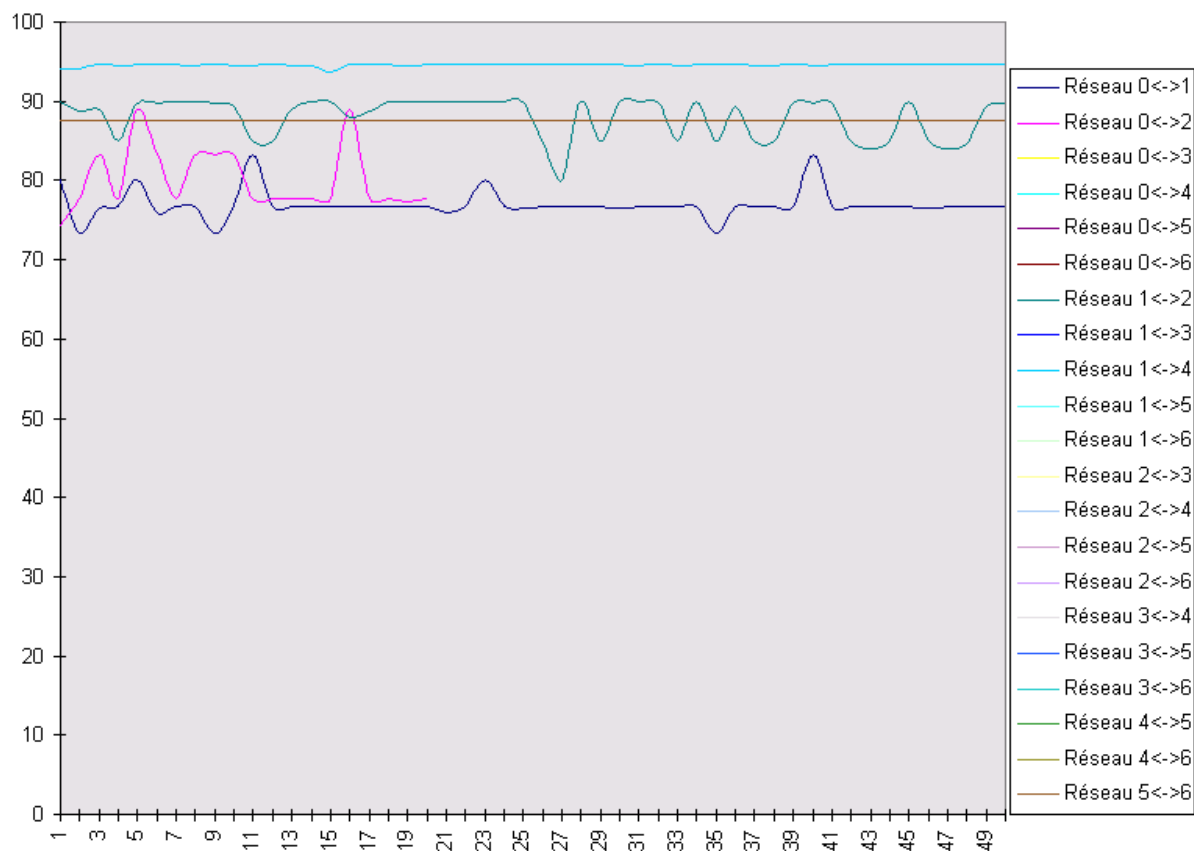


Figure 11. Moyenne de la variation du taux de reconnaissance (sur l'ensemble d'évaluation) en fonction du nombre de neurones de la couche cachée des réseaux liés à la base "glass".

1.8.5. La sélection par élimination

Nous allons tout d'abord comparer l'influence du critère d'ordonnement utilisé pour créer la hiérarchie. Le tableau 6 présente les taux de reconnaissance obtenus sur les bases de données (ensembles de tests) pour les deux critères. On constate qu'il y a peu de variation entre les deux critères même si le critère $|S(\mathcal{R}_k) \times Q(\mathcal{R}_k)|$ semble meilleur.

Base	$Q(\mathcal{R}_k)$	$ S(\mathcal{R}_k) \times Q(\mathcal{R}_k) $
Wine	100,0%	100,0%
Vehicle	78,4%	78,4%
PageBlocks	90,1%	90,1%
Glass	82,2%	82,2%
Segmentation	87,8%	87,7%
OptDigits	81,6%	83,2%
Serous	65,8%	66,3%

Tableau 6. Comparaison des taux de reconnaissance pour les deux critères de sélection par élimination.

Si l'on s'intéresse maintenant à la sélection par élimination, nous pouvons comparer les sélections numéro 1 et 2. Un exemple de classification par le critère $Q(\mathcal{R}_1)$ en utilisant les sélections 1 et 2 donne respectivement :

Classification par le critère $Q(\mathcal{R}_1)$ et la sélection numéro 1 :

réseau n°0 entre 0 et 3 qualité 100.00 sortie 0.023845
réseau n°1 entre 0 et 4 qualité 100.00 sortie 0.085131
réseau n°2 entre 0 et 5 qualité 100.00 sortie 0.150728
réseau n°3 entre 0 et 6 qualité 100.00 sortie 0.096272
réseau n°4 entre 1 et 3 qualité 100.00 sortie 0.022629
réseau n°5 entre 1 et 5 qualité 100.00 sortie 0.810577
réseau n°6 entre 2 et 3 qualité 100.00 sortie 0.016257
réseau n°7 entre 2 et 4 qualité 100.00 sortie 0.866629
réseau n°8 entre 2 et 5 qualité 100.00 sortie 0.372360
réseau n°9 entre 2 et 6 qualité 100.00 sortie 0.351300
réseau n°10 entre 3 et 4 qualité 100.00 sortie -0.014316
réseau n°11 entre 3 et 5 qualité 100.00 sortie -0.034243
réseau n°12 entre 3 et 6 qualité 100.00 sortie -0.009821
réseau n°13 entre 4 et 5 qualité 100.00 sortie -0.333740
réseau n°14 entre 4 et 6 qualité 100.00 sortie -0.076392
réseau n°15 entre 1 et 4 qualité 94.73 sortie 0.973088
réseau n°16 entre 1 et 6 qualité 90.90 sortie -0.000293
réseau n°17 entre 1 et 2 qualité 90.00 sortie -0.000285
réseau n°18 entre 0 et 2 qualité 88.88 sortie 0.003078

Classification par le critère $Q(\mathcal{R}_1)$ et la sélection numéro 2 :

réseau n°0 entre 0 et 3 qualité 100.00 sortie 0.023845
réseau n°1 entre 0 et 4 qualité 100.00 sortie 0.085131
réseau n°2 entre 0 et 5 qualité 100.00 sortie 0.150728
réseau n°3 entre 0 et 6 qualité 100.00 sortie 0.096272
réseau n°16 entre 1 et 6 qualité 90.90 sortie -0.000293
réseau n°18 entre 0 et 2 qualité 88.88 sortie 0.003078

On dispose de l'enchaînement des réseaux, de leur numéro dans la hiérarchie, des classes discriminées par un réseau, de la qualité de classification du réseau et de la valeur du neurone de sortie. Ces deux schémas de classification amènent à la même décision : la classe 0 a été choisie. Cependant, on remarque que le nombre de réseaux participant à la décision de classification est plus élevé pour la sélection numéro 1 (18 contre 6). Cela ne permet pas d'utiliser moins de réseaux car il faut de toute façon calculer la sortie d'un réseau pour savoir quelle classe éliminer, mais cela permet d'avoir une vision plus hiérarchique de la décision de classification.

Si l'on compare enfin les deux critères d'ordonnement, on constate que la classification réalisée par le critère $|S(\mathcal{R}_k) \times Q(\mathcal{R}_k)|$ utilise des réseaux différents mais qui amènent à la même décision de classification.

Classification par le critère $|S(\mathcal{R}_k) \times Q(\mathcal{R}_k)|$ et la sélection numéro 2 :

réseau n°15 entre 1 et 4 qualité 94.73 sortie 0.973088
 réseau n°5 entre 1 et 5 qualité 100.00 sortie 0.810577
 réseau n°9 entre 2 et 6 qualité 100.00 sortie 0.351300
 réseau n°11 entre 3 et 5 qualité 100.00 sortie -0.034243
 réseau n°20 entre 0 et 1 qualité 83.33 sortie 0.004212
 réseau n°18 entre 0 et 2 qualité 88.88 sortie 0.003078

Le tableau 7 présente pour chacune des bases, le nombre moyen de réseaux permettant de prendre une décision de classification. On constate que suivant la sélection par élimination employée, le nombre va croissant au fur et à mesure que le nombre de classes augmente. Mais la sélection numéro 2 permet de réduire cette augmentation de façon significative (tableau 7). La sélection à retenir est donc la sélection numéro 2 pour disposer d'un point de vue plus hiérarchique de la décision de classification.

Base	Classes	Sélection 1	Sélection 2
Wine	3	2	2
Vehicle	4	5	3
PageBlocks	5	8	4
Glass	7	19	6
Segmentation	7	13	6
OptDigits	9	31	9
Serous	18	130	17
Moyenne	7	30	7

Tableau 7. Le nombre moyen de réseaux de neurones associés à la décision de classification selon la sélection utilisée.

1.8.6. Comparaison générale

Pour terminer, nous allons comparer les résultats obtenus par notre architecture avec ceux d'autres travaux et plus précisément ceux de Richeldi [Richeldi 96] et de Yang [Yang 98a, 98b]. Ces travaux sont basés sur des réseaux de neurones. ADHOC [Richeldi 96] est un réseau de neurones classique utilisant une sélection de caractéristiques. DistAI [Yang 98a] utilise des algorithmes génétiques pour l'apprentissage, et GA-DistAI [Yang 98b] est une version améliorée de DistAI [Yang 98a] utilisant une sélection de caractéristiques, cette sélection faisant appel également à des algorithmes génétiques. On constate que, sur les bases de données communes à notre étude, notre architecture est meilleure sur toutes les bases présentant plus de 2 classes (tableau 8). Ce fait est accentué lorsque le nombre de classes augmente. Pour un nombre de classes égal à deux notre architecture ne se distingue pas de la structure d'un RMC classique, ce qui explique des résultats moindres par rapport à des RMC

utilisant des techniques d'amélioration (algorithmes génétiques et sélection de caractéristiques).

Base	MONNA	ADHOC	DistAI	GADistAI
Pima	76,6%	73,2%	76,3%	79,5%
Ionosphere	90,1%	-	94,3%	98,6%
Cancer	97,8%	-	97,8%	99,3%
Wine	100,0%	-	97,1%	99,4%
Vehicle	78,4%	69,6%	65,4%	68,8%
Glass	82,2%	70,5%	70,5%	80,8%

Tableau 8. Comparaison entre différentes approches neuronales.

2. Sélection d'attributs

Les réseaux utilisés par l'architecture MONNA sont généralement très simples, certains n'utilisent d'ailleurs qu'un seul neurone sur la couche cachée. L'algorithme de structure adaptative permet d'obtenir ce genre de structure très simple. Par ce biais nous avons pu simplifier chaque réseau en réduisant le nombre de neurones de la couche cachée. Cependant, nous pouvons simplifier à nouveau ces réseaux, car une des autres couches restantes peut être simplifiée : la couche d'entrée. En effet, en supprimant certaines des entrées et donc en éliminant les attributs non pertinents, on peut arriver à des résultats similaires voire meilleurs en terme de taux de reconnaissance. Ceci nous relie à un thème de recherche bien connu en classification : la sélection d'attributs. Dans ce paragraphe consacré à la sélection d'attributs, nous nous attacherons tout d'abord à définir la notion d'attribut pertinent. Nous traiterons le problème d'un point de vue général concernant la recherche heuristique, puis nous verrons les différentes approches qui ont été développées pour résoudre ce problème de sélection d'attributs, dénommées par *filter* et *wrapper*. Nous verrons quelques-unes des approches communément utilisées en classification.

2.1. Le problème de la sélection d'attributs

Un problème universel auquel sont confrontés les systèmes intelligents consiste à déterminer sur quoi porter leur attention. Ce genre de systèmes doit décider quels aspects du problème sont pertinents, un système expert doit, par exemple, déterminer quels attributs utiliser dans les règles. Étant donné que chaque processus de classification apprend à partir d'expériences, la distinction entre les attributs pertinents et non pertinents est un problème difficile. On consultera [Dash 97, Doak 92, Langley 94, Siedlecki 88] pour avoir un panorama complet des problèmes et méthodes en sélection d'attributs.

La majorité des problèmes de classification requièrent un apprentissage sur un ensemble de données dont chaque instance est associée à une classe ou un label. Chaque instance est décrite par un vecteur d'attributs (ou de caractéristiques) et une classe. La tâche d'un algorithme d'induction est de construire un classifieur qui permet de classer les différentes instances qui lui seront proposées. Le classifieur est donc une projection de l'espace des attributs vers l'ensemble des classes possibles. La tâche qui consiste à construire un classifieur (l'induction) qui permet de classer de manière optimale de nouvelles instances peut cependant être améliorée en sélectionnant les attributs qui servent d'entrée au classifieur [Fukunaga 90, Yang 98]. Les attributs qui sont présentés à un classifieur peuvent affecter :

- Le taux de reconnaissance du classifieur qui est obtenu en utilisant un algorithme d'apprentissage (arbre de décision ou réseaux de neurones) : les attributs utilisés pour décrire les instances de l'ensemble d'apprentissage définissent implicitement un langage, qui, s'il n'est pas assez expressif, ne permettra pas d'obtenir une généralisation suffisante. Cela aura pour effet de limiter le taux de reconnaissance du classifieur [Blum 97].

- Le temps nécessaire pour un apprentissage permettant d'obtenir un taux de reconnaissance satisfaisant : plus le nombre d'attributs est élevé, plus l'espace de recherche devant être exploré par l'algorithme d'apprentissage est élevé. Un nombre élevé d'attributs non pertinents augmentera la taille de l'espace de recherche et ainsi le temps d'apprentissage.

- Le nombre d'exemples nécessaires pour obtenir un taux de reconnaissance satisfaisant : plus le nombre d'attributs décrivant une instance est élevé, plus le nombre d'exemples nécessaires pour obtenir un taux de reconnaissance satisfaisant est élevé.

La nécessité de sélectionner les attributs devant servir à créer le classifieur à partir d'un algorithme d'induction s'impose donc. Cependant, nous ne savons pas encore comment définir la pertinence d'un attribut présenté à un classifieur.

2.2. Définition de la pertinence

Dans cette section, nous présentons les définitions de la pertinence d'attributs qui ont été suggérées dans la littérature. Un algorithme de classification dispose d'un ensemble de n données constituant les ensembles d'apprentissage et de test. Chaque instance X est un élément de l'ensemble $F_1 \times F_2 \times \dots \times F_m$, où F_i est le domaine de variation de l'attribut i . Les instances des ensembles d'apprentissage et de test sont des associations attributs-classe $\langle X, Y \rangle$ où Y est la classe de l'instance X . Pour chaque instance, on définit x_i comme la valeur de l'attribut X_i .

La tâche d'un algorithme de classification est de déterminer une structure (par exemple un réseau de neurones ou un arbre de décision) de telle façon que pour une nouvelle instance, il est possible de prédire de façon correcte la classe Y de cette instance, ceci définit une probabilité p dans l'espace $F_1 \times F_2 \times \dots \times F_m \times Y$. Nous noterons comme c le classifieur utilisé, on dispose alors de $c(X)$ comme étant la classe de l'instance X .

Almuallim et Dietterich [Almuallim 1991] définissent la pertinence d'un attribut en supposant que tous les attributs sont booléens et qu'il n'y a pas de bruit parmi les données :

Définition 1

Un attribut X_i est défini comme pertinent pour un classifieur c si X_i apparaît dans chaque formule booléenne qui représente c et est non pertinent dans les autres cas.

Gennari, Langley et Fisher [Gennari 1989] autorisent la présence de bruit dans les données ainsi que des attributs multi-valués. Il définissent la pertinence d'une variable par :

Définition 2

Un attribut X_i est défini comme pertinent s'il existe des x_i et y pour lesquels $p(X_i = x_i) > 0$ tels que $p(Y = y \mid X_i = x_i) \neq p(Y = y)$

Par cette définition, X_i est pertinent si le fait de connaître sa valeur permet de changer l'estimation de la classe Y , c'est-à-dire que Y est conditionnellement dépendant de X_i . Cette définition n'est pas correcte dans le cas où toutes les probabilités des instances non étiquetées sont équiprobables et la définition 2 doit être modifiée de la façon suivante :

Soit $S_i = \{X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_m\}$ l'ensemble de tous les attributs sauf X_i et soit s_i l'ensemble S_i dont tous les attributs ont été assignés à une valeur.

Définition 3

X_i est pertinent s'il existe x_i , y et s_i pour lesquels $p(X_i = x_i) > 0$ tel que

$$p(Y = y, S_i = s_i \mid X_i = x_i) \neq p(Y = y, S_i = s_i)$$

Par la définition 3, X_i est pertinent si la probabilité d'obtenir la classe y change lorsque l'on ne connaît pas la valeur de X_i .

Définition 4

X_i est pertinent s'il existe x_i, y et s_i pour lesquels $p(X_i = x_i, S_i = s_i) > 0$ tel que

$$p(Y = y \mid X_i = x_i, S_i = s_i) \neq p(Y = y \mid S_i = s_i)$$

Nous allons maintenant voir sur un exemple que ces définitions donnent des résultats incorrects. Cet exemple est tiré de l'article de John [John 94]. On définit un ensemble de 5 attributs booléens, X_1, \dots, X_5 . Les attributs X_2 et X_3 sont respectivement négativement corrélés avec X_4 et X_5 , c'est-à-dire que $X_4 = \overline{X_2}, X_5 = \overline{X_3}$. Il y a seulement 8 instances possibles dans l'espace des attributs $F_1 \times F_2 \times \dots \times F_5$ et les probabilités sont données comme équiprobables. Le classifieur c s'exprime de la façon suivante : $Y = X_1 \oplus X_2$ (\oplus signifie XOR).

On peut noter que l'on dispose d'un équivalent booléen à l'expression du classifieur : $Y = X_1 \oplus \overline{X_4}$. Les attributs X_3 et X_5 sont non pertinents au sens le plus large, X_1 est indispensable et l'on peut utiliser indépendamment X_2 ou X_4 . Le tableau donne pour chaque définition quels attributs sont pertinents et lesquels ne le sont pas.

Suivant la définition 1, X_3 et X_5 sont non pertinents ainsi que X_2 et X_4 car ils peuvent être remplacés par le complémentaire de l'autre. Suivant la définition 2, tous les attributs sont non pertinents car pour chaque classe y et pour chaque attribut x , il y a deux instances qui correspondent avec ces valeurs. Selon la définition 3, chaque attribut est non pertinent car la connaissance de la valeur de l'un d'entre eux change la probabilité de quatre des huit instances possibles de $1/8$ à 0 . Selon la définition 4, X_3 et X_5 sont non pertinents et X_2 et X_4 également car ils n'apportent aucune information à S_4 et S_2 .

Définition	Attributs pertinents	Attributs non pertinents
Définition 1	X_1	$X_2, X_3, X_4, X_5,$
Définition 2	Aucun	Tous
Définition 3	Tous	Aucun
Définition 4	X_1	$X_2, X_3, X_4, X_5,$

Tableau 9. La pertinence des attributs pour le problème XOR corrélé selon les quatre définitions de la pertinence d'un attribut.

Nous allons maintenant voir des notions de pertinence d'attributs qui sont utiles du point de vue d'un algorithme d'apprentissage qui doit déterminer quels attributs conserver et lesquels éliminer. Pour cela sont introduites deux notions très importantes : les attributs fortement et faiblement pertinents [Kohavi 94,95,97,98]. En effet, deux degrés de pertinence sont nécessaires. Un attribut X est fortement pertinent si la suppression de X entraîne une détérioration du taux de reconnaissance, un attribut X est faiblement pertinent s'il n'est pas fortement pertinent et s'il existe un sous-ensemble d'attributs, S , tel que le taux de reconnaissance du classifieur en utilisant S est plus élevé que le taux de reconnaissance du classifieur en utilisant $S \cup \{X\}$. Un attribut est non pertinent s'il n'est ni fortement, ni faiblement pertinent.

Définition 4 (forte pertinence)

X_i est fortement pertinent s'il existe x_i, y et s_i pour lesquels $p(X_i = x_i, S_i = s_i) > 0$ tel que

$$p(Y = y \mid X_i = x_i, S_i = s_i) \neq p(Y = y \mid S_i = s_i)$$

Définition 5 (Faible pertinence)

Un attribut X_i est faiblement pertinent si il n'est pas fortement pertinent et s'il existe un sous ensemble d'attributs S'_i de S_i pour lequel il existe x_i, y et s_i avec $p(X_i = x_i, S'_i = s'_i) > 0$ tel que

$$p(Y = y \mid X_i = x_i, S'_i = s'_i) \neq p(Y = y \mid S'_i = s'_i)$$

Dans l'exemple donné précédemment (le XOR corrélé), l'attribut X_1 est fortement pertinent, X_2 et X_4 sont faiblement pertinents et X_3 et X_5 sont non pertinents.

Les attributs qui sont fortement pertinents sont importants et doivent être conservés, simplement parce que l'élimination d'un attribut fortement pertinent fait apparaître des ambiguïtés parmi les données. Les attributs qui sont faiblement pertinents peuvent être ou ne pas être importants à conserver, cela dépend de la suppression éventuelle d'autres attributs.

Nous pouvons à présent définir le problème principal posé par la sélection d'attributs : trouver un sous-ensemble optimal selon un classifieur donné. Le problème de la sélection de caractéristiques se réduit donc à déterminer un sous-ensemble optimal.

Définition 6

Pour un classifieur c , et une base de données D possédant les caractéristiques X_1, X_2, \dots, X_m , un sous-ensemble optimal d'attributs S_{opt} est un sous-ensemble d'attributs tel que le taux de reconnaissance du classifieur soit maximal.

Un sous-ensemble optimal d'attributs n'est pas nécessairement unique. Il se peut qu'il soit possible d'obtenir le même taux de reconnaissance en utilisant différents sous-ensembles d'attributs, ceci est le cas lorsque deux attributs sont parfaitement corrélés et qu'ils peuvent être remplacés l'un par l'autre indépendamment.

2.3. Composantes de base des méthodes de sélection de caractéristiques

Selon Dash [Dash 97] la sélection d'attributs peut être vue sous différents points de vue, mais ils sont souvent similaires. Dash définit un ensemble de buts conceptuellement différents qui peuvent être atteint par un algorithme de sélection d'attributs.

1. *Idéal* : trouver le sous-ensemble minimal qui est nécessaire et suffisant pour le classifieur,

2. *Classique* : Sélectionner un sous-ensemble de m' attributs parmi un ensemble de m attributs ($m' < m$), tel que la valeur d'un critère donné soit optimale pour le sous-ensemble de taille m' ,
3. *Augmentation du taux de reconnaissance* : le but de la sélection d'attributs est de choisir un ensemble d'attributs qui augmente le taux de reconnaissance ou bien qui simplifie la structure du classifieur sans qu'il y ait de baisse du taux de reconnaissance initial (en utilisant tous les attributs),
4. *Approximation de la distribution des classes originales* : le but de la sélection de caractéristiques est de sélectionner le plus petit sous-ensemble pour lequel la distribution des classes est aussi proche que possible de la distribution originale utilisant tous les attributs.

La sélection des attributs sélectionne le sous-ensemble de taille minimale suivant la valeur d'un critère qui est parmi :

1. Le taux de reconnaissance ne décroît pas de façon significative,
2. La distribution des classes pour un sous-ensemble donné est aussi proche que possible de la distribution originale des classes utilisant tous les attributs.

On peut finalement synthétiser un algorithme de sélection d'attributs en plusieurs composantes de base (Figure 12) :

1. **Un algorithme de recherche d'un sous ensemble.** Il parcourt l'espace des sous-ensembles d'attributs, qui est de taille $2^m - 1$ où m est le nombre d'attributs.
2. **Une fonction d'évaluation.** Elle prend un sous-ensemble d'attributs en entrée et donne une valeur numérique en sortie. Elle sert à évaluer le sous-ensemble qui est examiné. Le but de l'algorithme de recherche est de maximiser cette fonction d'évaluation.
3. **Un inducteur.** C'est un algorithme qui permet de générer un classifieur à partir de données d'apprentissage. Le classifieur obtenu est utilisé pour classer les instances de la base de données.
4. **Un critère d'arrêt.** Ce critère permet de déterminer quand l'algorithme de recherche doit se terminer. Cela peut éventuellement être une stratégie de sélection de classifieur.
5. **Une procédure de validation.** Cela permet de déterminer si le sous-ensemble sélectionné par la sélection de caractéristiques est valide.

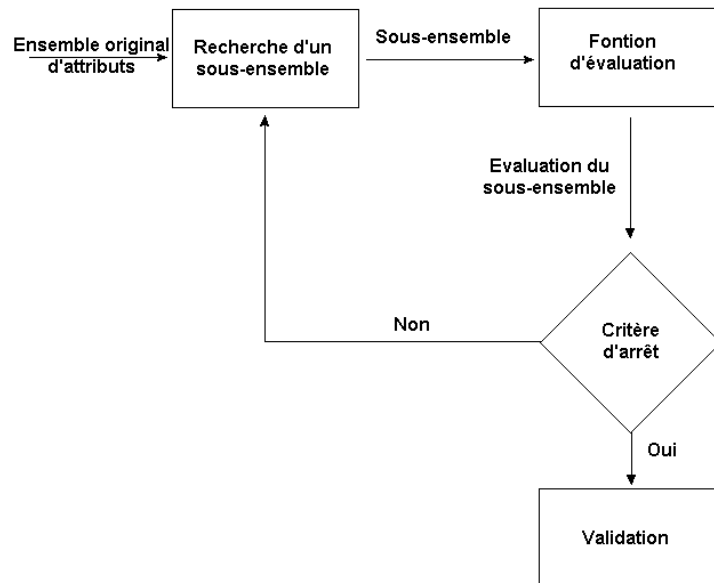


Figure 12. Les différentes composantes d'un procédé de sélection d'attributs.

L'algorithme de recherche d'un sous ensemble consiste à générer des sous-ensembles destinés à la fonction d'évaluation. Si l'ensemble original des attributs est de taille m , alors le nombre total de sous-ensembles possibles est $2^m - 1$. Ce nombre peut devenir rapidement élevé même pour des valeurs faibles de m . L'algorithme de génération peut commencer à partir d'un sous-ensemble qui est soit vide, soit qui contient tous les attributs, soit à partir d'un sous-ensemble obtenu de façon aléatoire. Le parcours de l'espace de recherche définit le choix des sous-ensembles.

La fonction d'évaluation mesure la pertinence d'un sous-ensemble produit par la procédure de recherche et de génération. La valeur donnée par l'évaluation d'un sous-ensemble est comparée avec la meilleure valeur précédemment obtenue. Lorsqu'un sous-ensemble plus performant (au sens de la fonction d'évaluation) est découvert, celui-ci est retenu comme étant le meilleur.

Le critère d'arrêt dépend généralement du choix des deux précédentes composantes (l'algorithme de recherche et la fonction d'évaluation). Ce critère peut être basé sur l'algorithme de recherche à travers le nombre d'attributs sélectionnés ou le nombre de sous-ensembles évalués (le nombre d'itérations). Il peut être également basé sur la fonction d'évaluation si les sous-ensembles générés ne produisent pas un meilleur sous-ensemble ou bien si un sous-ensemble généré a atteint une valeur maximale ou satisfaisante de la fonction d'évaluation.

La procédure de validation ne fait pas partie intégrante du procédé de sélection d'attributs mais cependant le sous-ensemble final doit être évalué afin de vérifier son adéquation aux objectifs. Cela consiste à effectuer différents tests comparant les résultats obtenus avec le sous-ensemble sélectionné et les précédents résultats utilisant tous les attributs.

Nous allons maintenant voir plus en détail les deux principales composantes d'un algorithme de sélection de caractéristiques : l'algorithme de recherche et de génération ainsi que la fonction d'évaluation.

2.3.1. Les espaces de recherche

La sélection d'attributs est essentiellement un problème d'optimisation qui demande l'exploration de l'espace constitué de tous les sous-ensembles afin d'identifier le sous-ensemble qui est optimal selon la fonction d'évaluation. On peut classer les algorithmes de sélection de variables selon trois catégories correspondant chacune à la stratégie employée pour parcourir l'espace de recherche exhaustive, heuristique et aléatoire [Doak 92].

- La recherche exhaustive

Ici, l'algorithme de recherche réalise un parcours complet de tout l'espace de recherche afin de trouver le sous-ensemble optimal. Différentes heuristiques peuvent être utilisées afin de réduire l'espace de recherche sans diminuer les chances de trouver le sous-ensemble optimal. Ces algorithmes garantissent le caractère optimal de la recherche car il peuvent revenir en arrière dans l'espace de recherche. On notera pour cela les algorithmes tels que le *branch and bound* [Narendra 77], le *best first search* (algorithme Focus) [Almuallim 94] et le *minimum description length* [Rissanen 78, Sheinvald 90]. On notera que ce genre de recherche est infaisable en pratique sauf si le nombre d'attributs est peu élevé.

- La recherche heuristique

Puisqu'une recherche exhaustive parmi tous les sous-ensembles d'un ensemble d'attributs est infaisable en pratique (quand le nombre d'attributs est élevé), de nombreux auteurs ont utilisé des algorithmes heuristiques. Il existe beaucoup de variations à ce procédé de recherche mais l'on peut dire que le principe de la génération de sous-ensembles se fait par ajout et/ou suppression. La complexité de ces algorithmes est polynomiale (en $O(m^2)$). Ces algorithmes se rapprochent souvent d'une recherche de type *branch and bound*. En représentant, pour un nombre d'attributs donné, l'espace de recherche complet pouvant être parcouru par un algorithme de recherche, on s'aperçoit qu'il y a une relation d'ordre car chaque sous-ensemble enfant a exactement un attribut de plus ou de moins que ses sous-ensembles parents (Figure 13). Cela suggère qu'une recherche heuristique doit soit commencer par un ensemble vide et ajouter des attributs, soit par l'ensemble de tous les attributs et les éliminer successivement. Ces deux approches s'appellent respectivement *forward sequential selection (FSS)* et *backward sequential selection (BSS)*. Ce point concerne seulement le point de départ de la recherche heuristique. La sélection par *forward* choisit le paramètre qui, ajouté au sous-ensemble précédent, permet d'obtenir une meilleure valeur de la fonction d'évaluation. La sélection par *backward* choisit le paramètre qui, supprimé au sous-ensemble précédent permet d'améliorer la valeur de la fonction d'évaluation. Ces deux techniques sont optimales à chaque étape de la génération d'un nouveau sous-ensemble mais sont incapables d'anticiper les interactions complexes présentes entre les attributs et qui peuvent affecter la performance du classifieur. Selon Aha [Aha 96], on ne peut pas dire quel algorithme (*BSS* ou *FSS*) donnera les meilleurs résultats et cela doit être systématiquement testé lors de l'utilisation de l'une ou de l'autre méthode. On trouve d'autres méthodes qui alternent des étapes de type *backward* et *forward* en alternant l'ajout et le retrait des attributs (méthodes de type *stepwise*) [Kittler 86].

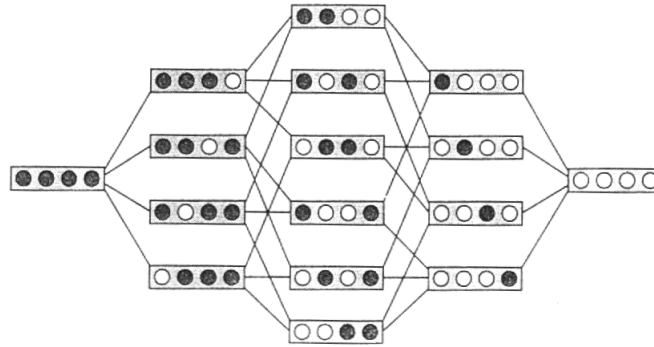


Figure 13. Chaque état de l'espace des attributs spécifie les attributs à utiliser. Chaque enfant (à droite) possède un attribut (ronds noirs) de moins que ses parents [Blum 97].

- La recherche aléatoire

Les algorithmes utilisant la recherche par des méthodes aléatoires sont relativement nouveaux. Bien que l'espace de recherche soit en $O(2^m)$, ces méthodes utilisent substantiellement moins de 2^m sous-ensembles en utilisant un nombre d'itérations maximal. Chaque génération aléatoire nécessite d'utiliser des paramètres dont la valeur doit être choisie avec soin. Ce genre d'algorithme de recherche fait appel à des procédés soit d'échantillonnage, soit aléatoires, soit probabilistes [Kohavi 94]. L'algorithme Relief [Kira 92] assigne des poids aux attributs en utilisant des instances obtenues aléatoirement. Les algorithmes *Las Vegas Filter* et *Las Vegas Wrapper* [Liu 96a, 96b] sont des algorithmes qui génèrent des sous-ensembles de façon aléatoire. Plusieurs auteurs se sont tournés vers l'utilisation d'algorithmes génétiques et ceci en utilisant des arbres de décision et des algorithmes de type K-Plus Proches Voisins [Hallinan 98, Richeldi 96, Siedlecki 89, Vafaie 93, Yang 98a, Yang 98b].

2.3.2. Les fonctions d'évaluation

Un sous-ensemble est donné comme optimal relativement à une fonction d'évaluation. Si bien qu'un autre sous-ensemble sera sûrement obtenu si l'on utilise une autre fonction d'évaluation. Le but d'une fonction d'évaluation est de mesurer la capacité de discrimination d'un attribut (et donc sa pertinence) ou d'un ensemble d'attributs. Dash [Dash 97] considère qu'il existe plusieurs sortes de fonction d'évaluation.

- La mesure de distance. C'est une mesure discriminatoire, ce qui nous relie directement à la mesure de la distance interclasse. Plusieurs distances provenant du domaine des statistiques peuvent être utilisées (Euclidienne, Fisher, Hausdorff [Jurie 99, Tomczak 99], etc.).
- La mesure d'information. Cela consiste à mesurer l'apport d'information d'un attribut. Un attribut est considéré meilleur qu'un autre si son apport d'information est plus élevé. Cette mesure se fait généralement par l'entropie [Quinlan 93].
- La mesure de la dépendance. Cela consiste à prédire la valeur d'un attribut à partir de la valeur d'un autre. Pour ce genre de fonction d'évaluation, on utilise la mesure de la corrélation entre un attribut et une classe.
- La mesure du taux de reconnaissance du classifieur. Cela consiste à utiliser le taux de reconnaissance du classifieur utilisé pour sélectionner un sous-ensemble. Le meilleur sous-ensemble est celui qui donne le meilleur taux de reconnaissance.

Pour les trois premières fonctions d'évaluation, on ne peut faire de supposition a priori sur la précision du sous-ensemble généré. Mais pour les fonctions d'évaluation utilisant le taux de reconnaissance du classifieur, la précision du sous-ensemble généré est très élevée car le sous-ensemble généré sera utilisé par ce même classifieur. On peut cependant dire que le temps d'exécution de méthodes utilisant le taux de reconnaissance sera plus élevé car il faut effectuer un apprentissage à chaque évaluation d'un sous-ensemble. Ce genre de fonction d'évaluation semble tout de même plus appropriée car cela permet de prendre en compte le modèle de classifieur utilisé.

2.4. Approches *filter* et *wrapper*

Suivant que l'algorithme de recherche et de génération des sous-ensembles utilise ou non le taux de reconnaissance du classifieur comme fonction d'évaluation, on peut répartir les algorithmes de sélection d'attributs en deux catégories. Si la sélection d'attributs est réalisée indépendamment du classifieur (i.e. l'algorithme d'apprentissage), cette approche est dénommée par le terme *filter* (car les attributs sont filtrés avant l'algorithme d'apprentissage du classifieur). Sinon cette approche est dénommée par le terme *wrapper* [John 94]. Comme nous l'avons vu précédemment, l'approche *filter* est plus rapide que l'approche *wrapper*. Mais son principal désavantage provient du fait que la sélection des attributs se fait de façon totalement indépendante du classifieur. L'approche *wrapper* nécessite, quand à elle, d'effectuer un apprentissage du classifieur pour chaque sous-ensemble à évaluer (Figure 14). Ceci est prohibitif et ne doit être fait que lorsque l'apprentissage du classifieur utilisé est relativement rapide.

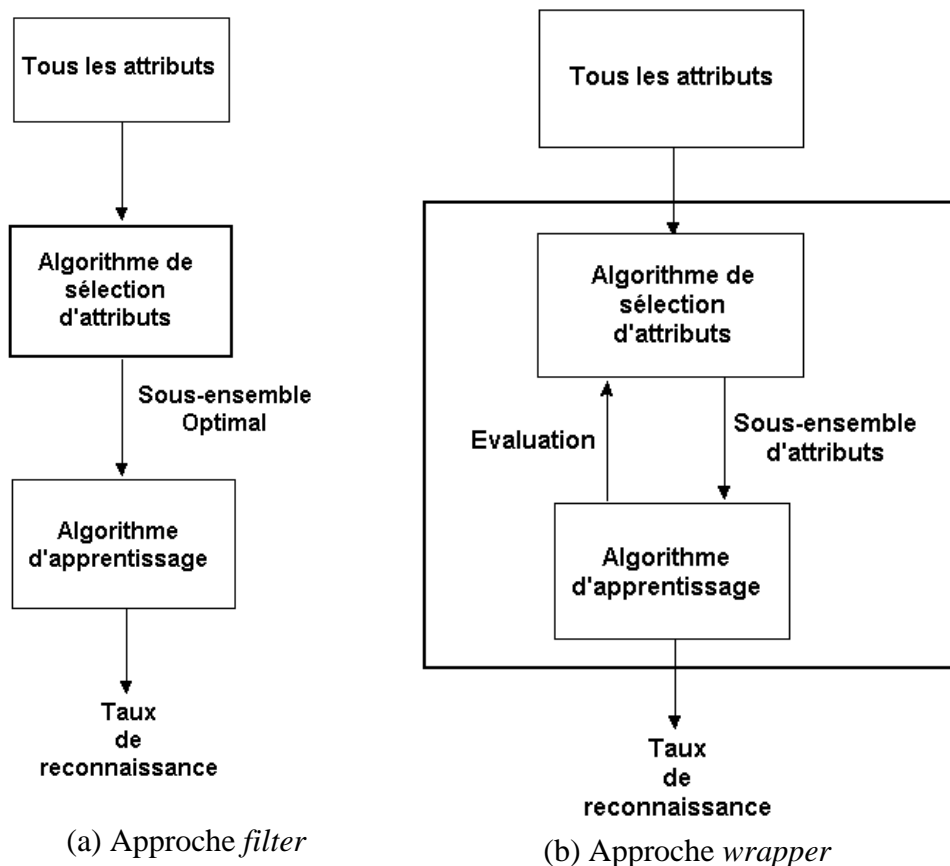


Figure 14. Les deux approches utilisées en sélection d'attributs, la différence se situe au niveau de l'incorporation du classifieur.

Seulement, il paraît préférable d'utiliser l'approche *wrapper* d'après Kohavi et Yang [Kohavi 94, Yang 98] pour la raison suivante. La plupart des mesures basées seulement sur les données (approche *filter*) que l'on trouve dans la littérature sont monotones. Cela signifie que pour une séquence de sous-ensembles obéissant à la relation $F_1 \supseteq F_2 \supseteq \dots \supseteq F_k$, la fonction d'évaluation (ou de mesure) f respecte le fait que $f(F_1) \geq f(F_2) \geq \dots \geq f(F_k)$. La monotonie de la fonction d'évaluation permet l'élagage de l'espace de recherche en utilisant un algorithme *branch and bound*. Mais la plupart des algorithmes d'apprentissage n'obéissent pas à cette restriction portant sur la monotonie de la fonction d'évaluation. Généralement, la plupart des algorithmes de sélection d'attributs qui utilisent l'approche *filter*, ignorent le fait que l'algorithme d'apprentissage n'est pas optimal. Les approches qui reposent sur la monotonie de la fonction d'évaluation (approches de type *filter*) semblent bien fonctionner avec des classifieurs linéaires mais donnent de mauvais résultats avec des classifieurs non linéaires tels que les réseaux de neurones.

Pour conclure, la sélection de caractéristiques ne doit pas être basée seulement en utilisant les données (approche *filter*), mais doit se faire relativement à un algorithme d'apprentissage. Ceci nous ramène à la définition 6 que nous avons donnée précédemment.

2.5. Algorithmes de recherche

Nous avons détaillé deux des composantes de base d'un algorithme de sélection de variables à savoir l'espace de recherche et la fonction d'évaluation. Il existe plusieurs façons de parcourir l'espace de recherche de manière heuristique (*FSS* ou *BSS*) et qu'il est préférable d'utiliser une approche de type *wrapper* pour chercher le sous-ensemble optimal. Nous allons voir dans ce paragraphe plusieurs algorithmes de recherche pouvant être utilisés avec l'approche *wrapper* (mais ils peuvent l'être également avec l'approche *filter*).

2.5.1. L'algorithme *hill-climbing*

La plus simple des techniques de recherche consiste à utiliser l'algorithme *hill-climbing* qui consiste à construire des sous-ensembles à partir d'un parent en prenant soin de toujours augmenter le taux de reconnaissance. L'algorithme se termine lorsque aucun sous-ensemble généré par un sous-ensemble parent n'apporte d'amélioration. On désigne par S l'ensemble initial des attributs.

1. Soit S l'ensemble initial,
2. Générer à partir de S par des opérateurs des sous-ensembles S' ,
3. Appliquer la fonction d'évaluation f aux sous-ensembles S' ,
4. Soit S'_{best} le meilleur sous-ensemble i.e. celui qui maximise $f(S')$,
5. Si $f(S') > f(S)$ alors $S = S'$ et aller à l'étape 2
6. S est le meilleur sous-ensemble.

Cet algorithme a tendance à se terminer sur des maxima locaux trop souvent, car dès que l'on constate une chute dans le taux de reconnaissance en appliquant un opérateur,

l'algorithme se termine. Les opérateurs pouvant être appliqués sont soit l'ajout soit le retrait d'un attribut.

2.5.2. L'algorithme *best-first search*

Cet algorithme est plus robuste que le précédent. L'idée consiste à sélectionner le meilleur sous-ensemble qui a été généré en vérifiant qu'il n'a pas déjà été généré. L'algorithme se termine lorsque l'on n'a pas réussi à trouver un meilleur sous-ensemble sur k générations successives. Un nœud est déclaré comme meilleur qu'un autre nœud s'il y a une augmentation de la valeur de la fonction d'évaluation.

On désigne par S l'ensemble initial des attributs, les opérateurs pouvant être appliqués à un sous-ensemble sont soit l'ajout soit le retrait d'un attribut. $Afaire$ désigne les sous-ensembles sur lesquels on applique les opérateurs et $Fait$ désigne les sous-ensembles déjà évalués.

1. $Afaire \leftarrow \{S\}$
 $Fait \leftarrow \{\emptyset\}$
 $Meilleur = S$
2. $S' = \arg \max_{E \in Afaire} f(E)$
3. $Afaire = Afaire - \{S'\}$
 $Fait = Fait \cup \{S'\}$
4. Si $f(S') > f(Meilleur)$ alors $Meilleur = S'$
5. Appliquer des opérateurs à S' donnant ses sous-ensembles enfants S'_i
6. Pour chaque $S'_i \notin Fait$ et $S'_i \notin Afaire$ évaluer S'_i et
 $Afaire = Afaire \cup \{S'_i\}$
7. Si $Meilleur$ a changé lors des k dernières itérations alors aller en 2
8. retourner $Meilleur$

2.5.3. L'algorithme *Sequential forward floating selection (SFFS)*

Cet algorithme a été proposé par Pudil [Pudil 94] et est une variante des différents algorithmes utilisant les opérateurs d'ajout et de retrait aux sous-ensembles générés. Selon Pudil, les algorithmes de type *SFS* ou *SBS* présentent des limitations. Le premier (*SFS*) ne peut plus éliminer un attribut qui a été sélectionné lors des étapes précédentes, de même l'algorithme *SBS* ne peut plus rajouter un attribut qui a été éliminé lors des étapes précédentes. Pour remédier à cela, Pudil propose d'utiliser non pas que des ajouts et que des suppressions d'attributs, mais les deux à la fois. L'algorithme *SFFS* est similaire à *SFS*, en ce sens que l'on part d'un ensemble d'attributs vide et que l'on ajoute progressivement des attributs. Cependant, après chaque ajout, *SFFS* tente par une série de suppressions successives d'éliminer les attributs faiblement pertinents de l'ensemble qui avait été généré. L'algorithme procède donc par ajouts et retraits successifs jusqu'à trouver le meilleur sous-ensemble possible. On désigne par S l'ensemble initial des attributs et par X_i l'attribut numéro i .

$S = \emptyset$
 $Meilleur = S$
 1. Si $f(S) > f(Meilleur)$ alors $Meilleur = S$
 2. Trouver X_i tel que $S' = \arg \max f(S \cup \{X_i\})$
 3. $S = S'$
 4. Trouver X_j tel que $S'' = S - \{X_j\}$ et $f(S'') > f(S)$
 5. Si $\exists S''$ tel que $f(S'') > f(S)$ aller en 1
 6. Sinon $S = S''$ et aller en 4
 7. retourner Meilleur

Cette méthode a été dénommée par le terme «float» car l'algorithme «flotte» de sous-ensembles en sous-ensembles en éliminant et ajoutant des attributs successivement.

2.6. Approches testées

Nous allons maintenant appliquer ces différentes approches de sélection de caractéristiques à l'architecture MONNA. Nous avons testé les approches *filter* et *wrapper*. Lors du paragraphe 3.4, nous avons précisé qu'une approche de style wrapper semble préférable pour effectuer une sélection de caractéristiques lorsque les classifieurs utilisés sont non-linéaires. Cependant, nous allons tester les deux approches à travers différents algorithmes, afin de comparer leur efficacité sur différentes bases.

2.7. L'approche filter

2.7.1. Trace des matrices de covariance inter et intra classe

Cet algorithme de sélection est basé sur l'approche *filter*, le meilleur sous-ensemble est donc fourni à l'inducteur avant tout apprentissage. Ici la fonction d'évaluation utilisée repose sur un critère de distance basé sur les matrices inter et intra classes.

Soit E un ensemble d'individus possédant une partition en k sous ensembles E_1, E_2, \dots, E_k . Un sous ensemble sera considéré comme un nuage de points. Les centres de gravité, inertie et matrice de covariance des sous ensembles E_i sont reliés par des relations qui jouent un rôle important en classification automatique.

- Centre de gravité des nuages

A partir de la partition de E en k classes, on a

$$E = \left\{ \overbrace{e_1, e_2, \dots, e_i}^{E_1}, \overbrace{\dots, e_l}^{E_2}, \dots, \overbrace{\dots, e_n}^{E_k} \right\}$$

$$1 = \underbrace{p_1 + p_2 + \dots + p_i}_{q_1} + \underbrace{\dots + p_l}_{q_2} + \dots + \underbrace{\dots + p_n}_{q_k}$$

On pose $q_i = \sum_{e_i \in E_i} p_i$ et on définit les centres de gravité des nuages E_j par $g_j = \frac{1}{q_j} \sum_i p_i e_i$ pour $e_i \in E$ et de l'ensemble E par $g = \sum_{j=1}^k g_j q_j$.

- Matrice de covariance inter et intra classe

La matrice $V_j = \frac{1}{q_j} \sum_{e_i \in E_j} p_i (e_i - g_j)^t \cdot (e_i - g_j)$ est la matrice de covariance de E_j . Par définition, on appelle W (comme *within*) la matrice intra classe définie comme la moyenne pondérée des V_j : $W = \sum_{j=1}^k q_j V_j$. De même, on définit la matrice interclasse B (comme *between*) comme la matrice de covariance des centres de gravité g_j affectés de leur masse q_j : $B = \sum_{j=1}^k q_j (g_j - g)^t \cdot (g_j - g)$. Cette matrice carré n'est pas inversible car les centres de gravité sont dans un sous-espace de dimension $k-1 < p$. On montre facilement que $V=W+B$.

- Inertie interclasse et intra classe

Le théorème de Koenig-Huygens [Diday 82] établit la relation entre l'inertie totale et l'inertie de chaque classe par rapport à son centre de gravité :

$I(E) = \sum_{j=1}^k I(E_j) + \sum_{j=1}^k q_j d_M(g_j, g)$ où $d_M(a, b) = (a - b)^t \cdot M \cdot (a - b)$ qui est la distance euclidienne si $M=I$.

L'inertie intra classe W est un critère d'évaluation d'une partition qui mesure l'homogénéité des classes de la partition. On cherche donc à minimiser ce critère. Par contre, on cherchera à maximiser B qui mesure la non homogénéité entre les classes de la partition. Pour mesurer l'efficacité d'un sous-ensemble de paramètres, nous allons utiliser le taux de traces des matrices interclasse et intra classe : $\beta = \frac{\text{Trace}(B)}{\text{Trace}(W)}$. Un sous-ensemble de paramètres est considéré comme le plus performant quand on obtient la plus petite valeur pour la matrice intra classe, ce qui revient à chercher à maximiser β .

L'approche de type *filter* que nous avons mise en œuvre, consiste à générer, par une approche de style *SFS*, des sous-ensembles et à retenir celui qui maximise β , nous donnant ainsi le meilleur sous-ensemble.

2.7.2. Théorie de l'information

Nous allons ici utiliser les algorithmes proposés par Quinlan [Quinlan 93] qui utilisent des fonctions d'entropie fondées sur la théorie de l'information. Reprenons E notre ensemble d'individus possédant une partition en k classes E_1, E_2, \dots, E_k , la quantité d'information obtenue

par une partition de E en k classes est donnée par $I(E) = - \sum_{j=1}^k \frac{|E_j|}{|E|} \times \log \left(\frac{|E_j|}{|E|} \right)$. Avec $|E_j|$

représentant le cardinal d'un ensemble E_j . Maintenant si l'on mesure la quantité d'information

apportée par le partitionnement de l'ensemble E en l sous-ensembles grâce à un attribut X_i , la quantité d'information de E par ce partitionnement s'écrit : $I_{X_i}(E) = \sum_{m=1}^l \frac{|E_m|}{|E|} \times I(E_m)$. On peut ensuite mesurer le gain d'information apporté par l'attribut X_i par $G(X_i) = I(E) - I_{X_i}(E)$. Un attribut sera plus informant qu'un autre (au sens de sa pertinence) si sa valeur du gain est plus élevée. Cependant la valeur du gain définie ci-dessus dépend du nombre de valeurs que peut prendre l'attribut X_i , en effet $G(X_i)$ tend à être élevé si la valeur de X_i est élevée. Pour pallier ceci, il faut utiliser une version normalisée du gain : $G'(X_i) = \frac{G(X_i)}{I'(X_i)}$ où I' est une fonction

de normalisation donnée par $I'(X_i) = \sum_{i=1}^n \frac{|E_i|}{|E|} \times \log\left(\frac{|E_i|}{|E|}\right)$.

L'algorithme utilisé est de type *filter* par une approche *SFS*, on sélectionne à chaque itération l'attribut qui donne la meilleure valeur du gain normalisé.

2.7.3. Méthode de Yacoub

La sélection d'attributs proposée par Yacoub [Yacoub 97] est spécifique aux réseaux de neurones. Cette méthode utilise seulement les paramètres d'un réseau à travers les poids et

la structure du réseau. Pour cela Yacoub utilise le critère $S_i = \sum_{j \in H} \left(\frac{|W_{ji}|}{\sum_{i \in I} |W_{ji}|} \sum_{k \in O} \frac{|W_{kj}|}{\sum_{j \in H} |W_{kj}|} \right)$ où I , H ,

O sont respectivement les couches d'entrée, cachée et de sortie. Cela permet de mesurer l'importance de l'attribut d'entrée i pour tous les neurones de sortie. L'algorithme utilise une recherche de type *backward* (*SBS*) qui s'arrête lorsque la performance du réseau décroît.

2.8. L'approche wrapper

Les algorithmes classiques que nous avons utilisé sont le *best-first search* en *SFS* et *SBS* mais également l'algorithme *SFFS*.

2.9. Résultats expérimentaux

Afin de quantifier l'apport de chacune des méthodes de sélection d'attributs, nous allons les comparer sur les bases de données du paragraphe 1.8.1. La sélection de caractéristiques est appliquée pour chacune des bases et différentes mesures concernant le taux de reconnaissance et le nombre d'attributs sélectionnés sont calculées. Pour chacune des approches, nous avons utilisé la règle suivante : si pour un réseau, pour lequel on sélectionne des attributs, le taux de reconnaissance en utilisant tous les attributs est meilleur que le taux de reconnaissance en utilisant le meilleur sous-ensemble, alors tous les attributs sont sélectionnés.

On notera que pour l'architecture MONNA, chaque réseau de neurones procède à une sélection d'attributs, contrairement à un réseau de neurones global qui sélectionne les attributs

pertinents pour séparer toutes les classes. Ici chaque réseau détermine les attributs pertinents pour différencier les deux classes qu'il sépare. Ceci a pour effet de simplifier l'architecture du réseau (au niveau de la couche d'entrée) mais également d'accroître le taux de reconnaissance de chacun des réseaux employés par l'architecture MONNA, ce qui peut permettre d'accroître le taux de reconnaissance global de l'architecture.

Nous allons tout d'abord comparer les méthodes de sélection d'attributs selon les taux de reconnaissance obtenus sur les bases de données. Puis nous comparerons les approches *filter* entre elles au niveau des attributs sélectionnés ainsi que pour les approches *wrapper*. Puis nous comparerons nos résultats avec ceux de la littérature des réseaux de neurones.

2.9.1. Comparaison sur les taux de reconnaissance

Nous allons tout d'abord comparer les différentes méthodes de sélection d'attributs que nous avons présentées : *SFS*, *SBS* (pour le *best-first search*), *SFFS*, *Yacoub*, *Cov* (matrice de covariance intra et inter classe) et *Quinlan*. Nous avons appliqué chacune de ces méthodes sur les bases de données du paragraphe 1.8.1. Le tableau 10 résume ces résultats. Les quatre premières bases présentent un nombre de classes à différencier égal à 2. Pour ces bases, l'architecture MONNA est constituée d'un seul RMC et la sélection d'attribut ne se fait que pour un seul réseau. Par contre pour les autres bases, le nombre de classes à différencier est supérieur à 2 si bien que la sélection d'attributs se fait pour plusieurs réseaux de neurones.

Sur les bases à deux classes la méthode *SFFS* donne toujours les meilleurs résultats et ceci quelle que soit la base. Sur la base Pima, on remarque que *SFS* et *SBS* donnent de moins bons résultats que toutes les autres approches et le taux de reconnaissance global n'a pas été amélioré. Or l'algorithme *SFFS* parvient à obtenir le même taux de reconnaissance que les autres meilleures méthodes (*Yacoub* et *Cov*), cela provient tout simplement du fait que pour *SFS* ou *SBS*, une fois qu'un attribut a été ajouté ou enlevé respectivement pour chacune des deux méthodes, aucun retour en arrière n'est possible. *SFFS* permet d'éviter ceci en procédant à des ajouts et des retraits successifs.

Base	MONNA	Yacoub	Cov	Quinlan	SBS	SFS	SFFS
Bupa	71,01%	71,01%	71,01%	71,01%	73,91%	73,91%	73,91%
Pima	76,62%	77,92%	77,92%	76,62%	76,62%	76,62%	77,92%
Ionosphere	90,14%	91,54%	90,14%	94,36%	92,95%	94,36%	95,77%
Cancer	97,81%	97,08%	97,81%	97,81%	98,50%	98,50%	98,54%
Wine	100%	100%	100%	100%	100%	100%	100%
Vehicle	78,36%	78,36%	79,53%	78,94%	84,21%	80,11%	82,45%
PageBlocks	90,05%	86,77%	92,79%	89,14%	94,71%	93,79%	94,98%
Glass	82,20%	84,44%	82,20%	80,00%	84,44%	84,44%	84,44%
Segmentation	87,81%	68,52%	82,43%	74,71%	80,42%	73,43%	72,66%
OptDigits	81,63%	87,31%	87,92%	86,02%	86,81%	84,36%	84,03%
Serous	65,80%	67,14%	71,98%	67,86%	74,97%	77,14%	77,14%

Tableau 10. Comparaison des différentes méthodes de sélection d'attributs employées.

Sur les bases présentant plus de 2 classes à différencier, la sélection d'attributs permet dans presque tous les cas d'augmenter le taux de reconnaissance par rapport à l'architecture MONNA sans sélection d'attributs. Seul le taux de reconnaissance sur la base Segmentation n'a pas été amélioré après sélection des attributs pertinents. On peut constater que lorsqu'il y a amélioration du taux de reconnaissance (Figure 15), globalement les méthodes de type

wrapper (*SBS*, *SFS*, *SFFS*) semblent meilleures que celles de type *filter* (Yacoub, Cov, Quinlan). Ceci n'est pas vérifié pour la base OptDigits pour laquelle la méthode Cov donne les meilleurs résultats.

On ne peut donc pas a priori établir de supériorité établie entre les méthodes *wrapper* et *filter*. Cependant, on peut faire certaines remarques. Dans tous les cas on peut dire que la sélection d'attributs en utilisant soit la méthode Cov, soit la méthode SFFS, permet d'améliorer les performances de l'architecture MONNA. Dans un cas la méthode Cov tient compte uniquement de la dispersion des données de la base et elle est donc bien adaptée dans le cas où les données permettent de vérifier la monotonie de la fonction d'évaluation, dans les autres cas la méthode SFFS doit être utilisée. Comme on ne peut pas vérifier cette monotonie de la fonction d'évaluation sans générer les sous-ensembles d'attributs, le plus simple est d'effectuer deux sélections d'attributs en utilisant ces deux méthodes et de sélectionner la meilleure. Cette approche semble très heuristique, mais on ne peut pas prédire *a priori* quelle méthode donnera les meilleurs résultats.

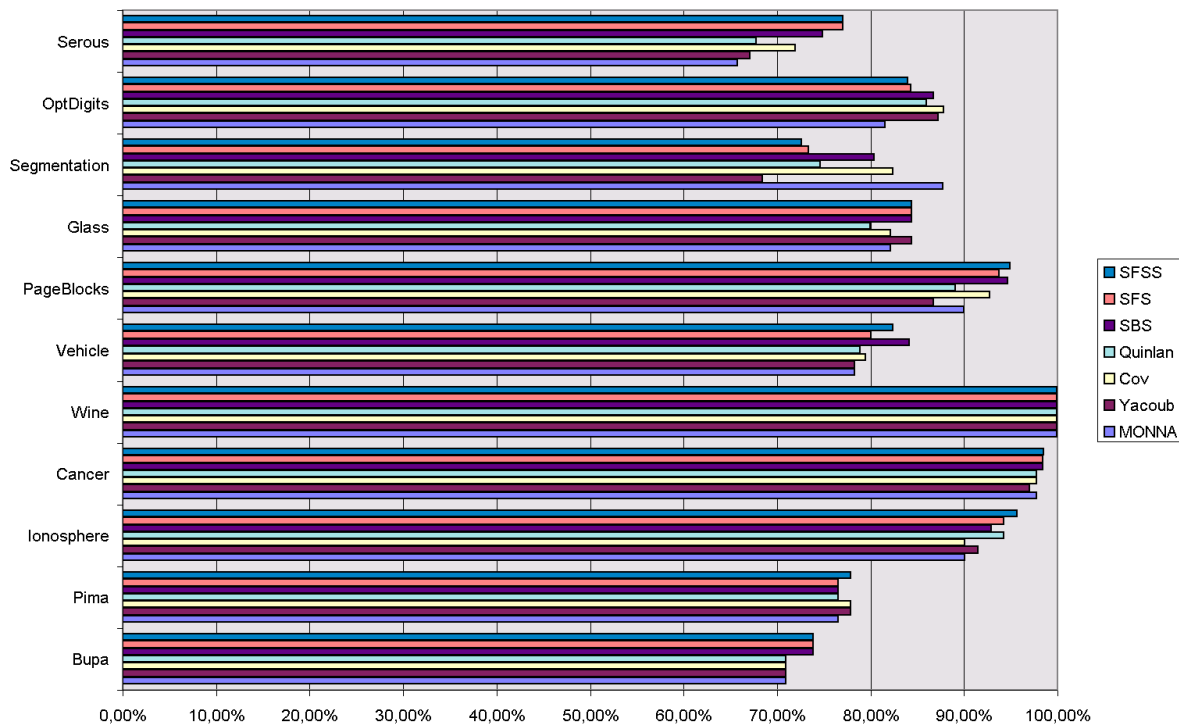


Figure 15. Histogramme comparant les différentes approches de sélection d'attributs.

2.9.2. Comparaison des approches *filter*

Nous allons à présent comparer les approches de type *filter* sur plusieurs points : le taux de reconnaissance obtenu mais également le nombre d'attributs sélectionnés pour les réseaux de l'architecture MONNA. Pour chacune des méthodes, on dispose du taux de reconnaissance de l'architecture MONNA après sélection des attributs, des nombres minimum, maximum et moyen d'attributs employés par chaque réseau ainsi que de leur écart type. Le tableau 11 rappelle le nombre de classes et d'attributs pour chacune des bases.

Base	Nombre de classes	Nombre d'attributs
Liver Bupa	2	6
Pima	2	8
Ionosphere	2	34
Cancer	2	9
Wine	3	13
Vehicle	4	18
PageBlocks	5	10
Glass	7	9
Segmentation	7	19
OptDigits	9	64
Serous	18	46

Tableau 11. Nombre de classes et d'attributs pour chaque classe servant aux expérimentations.

Base	Yacoub	Min	Max	Moy	Ecart-Type	Utilisés
Bupa	71,01%	5	5	5	0	5
Pima	77,92%	4	4	4	0	4
Ionosphere	91,54%	3	3	3	0	3
Cancer	97,08%	9	9	9	0	9
Wine	100%	8	12	9,66	1,69	13
Vehicle	78,36%	13	18	16,5	1,89	18
PageBlocks	86,77%	5	10	7,9	1,57	10
Glass	84,44%	2	9	6,9	1,92	9
Segmentation	68,52%	8	19	12,09	2,63	19
OptDigits	87,31%	58	64	62,53	1,31	64
Serous	67,14%	31	46	18,51	21,36	46

Tableau 12. Taux de reconnaissance et nombre d'attributs sélectionnés pour la méthode Yacoub sur les bases.

Base	Cov	Min	Max	Moy	Ecart-Type	Utilisés
Bupa	71,01%	6	6	6	0	6
Pima	77,92%	4	4	4	0	4
Ionosphere	90,14%	29	29	29	0	29
Cancer	97,81%	7	7	7	0	7
Wine	100%	3	13	6,66	4,49	13
Vehicle	79,53%	15	18	17	1,15	18
PageBlocks	92,79%	1	10	4	2,45	10
Glass	82,20%	2	9	5,47	3,01	9
Segmentation	82,43%	2	19	3,85	3,74	19
OptDigits	87,92%	2	63	39,08	16,35	64
Serous	71,98%	1	46	9,45	15,08	46

Tableau 13. Taux de reconnaissance et nombre d'attributs sélectionnés pour la méthode Cov sur les bases.

Base	Quinlan	Min	Max	Moy	Ecart-Type	Utilisés
Bupa	71,01%	6	6	6	0	6
Pima	76,62%	8	8	8	0	8
Ionosphere	94,36%	4	4	4	0	4
Cancer	97,81%	6	6	6	0	6
Wine	100%	4	11	7,33	2,86	11
Vehicle	78,94%	5	18	12	5,71	18
PageBlocks	89,14%	1	10	5,33	3,97	10
Glass	80,00%	3	9	6,38	2,25	9
Segmentation	74,71%	2	19	6,9	5,04	19
OptDigits	86,02%	22	64	45,31	10,85	64
Serous	67,86%	1	46	12,44	16,82	46

Tableau 14. Taux de reconnaissance et nombre d'attributs sélectionnés pour la méthode Quinlan sur les bases.

On dispose également du nombre total d'attributs utilisés par tous les réseaux. Ceci va nous permettre de comparer ces trois méthodes de sélection à plusieurs niveaux : le taux de reconnaissance mais également le nombre d'attributs employés et utilisés. Un attribut est dit employé quand il sert dans un réseau de l'architecture et un attribut est dit utilisé s'il sert dans l'architecture (il sert dans au moins un des réseaux de l'architecture).

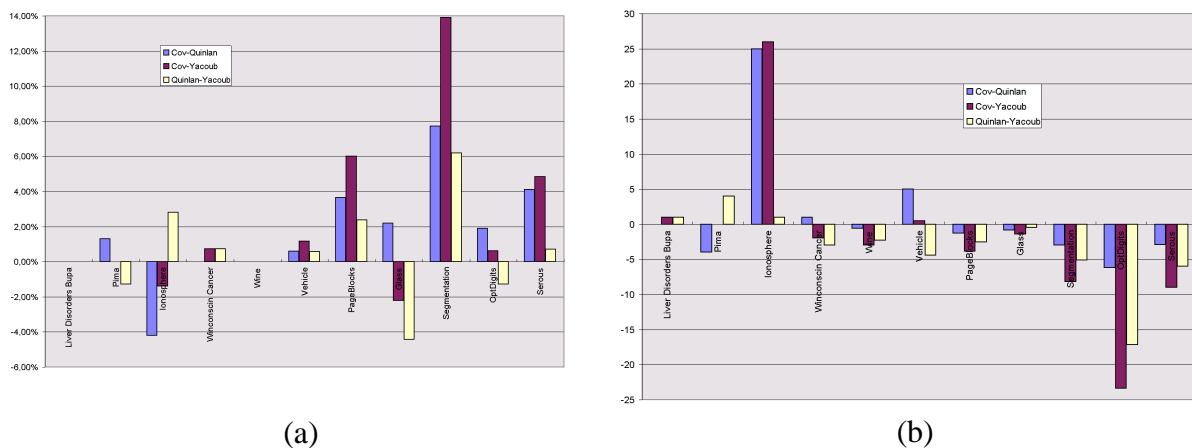


Figure 16. Ecarts entre les différentes méthodes utilisant l'approche *filter* pour le taux de reconnaissance global de l'architecture (a) et le nombre d'attributs moyen employés par chaque réseau (b).

Si l'on compare tout d'abord le taux de reconnaissance obtenu pour chacune des méthodes (Tableaux 12, 13, et 14 et figure 16), on constate que la méthode Cov est pratiquement toujours la meilleure sauf pour la base Glass (7 classes) et la base Ionosphere (2 classes). On note d'ailleurs que pour cette dernière aucune sélection d'attributs n'a été faite.

Le défaut majeur de la méthode de Yacoub est d'éliminer peu d'attributs (l'écart-type est peu élevé), ce qui semble handicapant pour une méthode de sélection de caractéristiques.

On remarque également que même après sélection, les attributs sont tous utilisés c'est-à-dire qu'ils sont employés dans au moins un réseau. Chacune des caractéristiques permet donc d'augmenter la pertinence d'un sous-ensemble, ceci se vérifie d'ailleurs car pour les

réseaux confrontés à une séparation difficile, le nombre d'attributs employés est généralement égal au nombre d'attributs utilisés.

2.9.3. Comparaison des approches *wrapper* sur le nombre d'attributs

Pour comparer les méthodes utilisant une approche de type *wrapper*, nous allons procéder de la même manière que pour l'approche *filter*. Les tableaux 15, 16, 17 et la figure 17. On constate tout d'abord que ce sont toujours les méthodes *SFFS* et *SBS* qui donnent les meilleurs taux de reconnaissance. La méthode *SFS* se comporte globalement de la même manière que la méthode *SFFS*. Ceci provient évidemment du fait que les deux algorithmes sont très proches, l'algorithme *SFFS* est meilleur car il permet à la fois d'ajouter et de retirer un attribut.

La méthode *SBS*, par contre, a tendance à entraîner un nombre moyen d'attributs employés élevé. Ceci provient de l'algorithme en lui-même car, dès qu'une chute apparaît dans le taux de reconnaissance, l'algorithme s'arrête. Or ceci peut être fréquent car on part de l'ensemble de tous les attributs : le taux de reconnaissance peut ne pas croître après avoir retiré un attribut pertinent.

Base	<i>SFS</i>	Min	Max	Moy	Ecart-Type	Utilisés
Liver	73,91%	4	4	4	0	4
Pima	76,62%	8	8	8	0	8
Ionosphere	94,36%	13	13	13	0	13
Cancer	98,50%	6	6	6	0	6
Wine	100%	2	3	2,33	0,47	6
Vehicle	80,11%	2	18	10,33	4,78	18
PageBlocks	93,79%	1	4	2,7	0,9	10
Glass	84,44%	2	9	4,71	2,86	9
Segmentation	73,43%	2	19	3,43	3,55	19
OptDigits	84,36%	3	33	11,02	7,46	61
Serous	77,14%	2	30	4,18	6,28	46

Tableau 15. Taux de reconnaissance et nombre d'attributs sélectionnés pour la méthode *SFS* sur les bases.

Base	<i>SBS</i>	Min	Max	Moy	Ecart-Type	Utilisés
Liver	73,91%	4	4	4	0	4
Pima	76,62%	8	8	8	0	8
Ionosphere	92,95%	29	29	29	0	29
Cancer	98,50%	4	4	4	0	4
Wine	100%	2	5	3,33	1,24	6
Vehicle	84,21%	4	18	10	4,65	18
PageBlocks	94,71%	1	8	4,5	2,15	9
Glass	84,44%	2	9	4,33	2,98	9
Segmentation	80,42%	2	4	2,09	0,42	5
OptDigits	86,81%	2	61	31,73	20,2	64
Serous	74,97%	2	45	8,71	13,28	46

Tableau 16. Taux de reconnaissance et nombre d'attributs sélectionnés pour la méthode *SBS* sur les bases.

Base	SFFS	Min	Max	Moy	Ecart-Type	Utilisé
Liver	73,91%	4	4	4	0	4
Pima	77,92%	4	4	4	0	4
Ionosphere	95,77%	20	20	20	0	20
Cancer	98,54%	6	6	6	0	6
Wine	100%	2	3	2,33	0,47	6
Vehicle	82,45%	5	12	7,5	2,36	17
PageBlocks	94,98%	2	8	4	2	10
Glass	84,44%	2	9	4,23	3,03	9
Segmentation	72,66%	2	19	2,95	3,61	19
OptDigits	84,03%	3	28	10,04	7,15	60
Serous	77,14%	2	30	4,18	6,28	46

Tableau 17. Taux de reconnaissance et nombre d'attributs sélectionnés pour la méthode SFFS sur les bases.

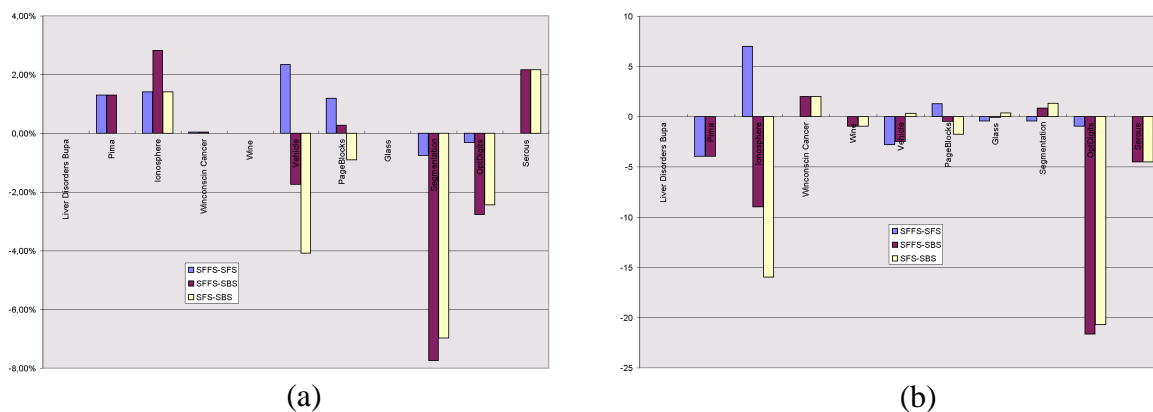


Figure 17. Ecart entre les différentes méthodes utilisant l'approche wrapper pour le taux de reconnaissance global de l'architecture (a) et le nombre d'attributs moyen employé par chaque réseau (b).

Mais le point le plus intéressant est que, contrairement aux approches *filter*, les approches *wrapper* permettent dans certains cas de sélectionner un nombre d'attributs utilisés qui est inférieur au nombre d'attributs initial. Ceci est très intéressant, car si un attribut n'est pas utilisé, il peut être totalement éliminé. Ceci entraîne l'augmentation du taux de reconnaissance et un gain dans le temps de calcul des caractéristiques.

3. Conclusion

Nous avons proposé une architecture de réseaux de neurones (MONNA) qui présente un intérêt pour les problèmes de classification présentant un grand nombre de classes. Puis nous avons appliqué plusieurs méthodes de sélection d'attributs (provenant des approches *filter* ou *wrapper*). Les algorithmes SFFS en *wrapper* et Cov en *filter* se sont avérés produire les meilleurs gains au niveau du taux de reconnaissance. La courbe de la figure 18 résume ces résultats et permet de constater le gain significatif obtenu par l'architecture MONNA sans et avec sélection de caractéristiques (Figure 18).

Pour terminer ces comparaisons, nous allons comparer le taux de reconnaissance de l'architecture MONNA avec et sans sélection d'attributs avec ceux obtenus dans d'autres travaux et plus précisément ceux de Richeldi [Richeldi 96] et de Yang [Yang 98a, 98b]. On rappelle que ADHOC [Richeldi 96] est un réseau de neurones utilisant une sélection d'attributs. DistAI [Yang 98a] utilise des algorithmes génétiques pour l'apprentissage et GA-DistAI [Yang 98b] est une version améliorée de DistAI utilisant une sélection d'attributs par algorithmes génétiques. On constate que, sur ces bases de données communes à notre étude, notre architecture est meilleure sur toutes les bases présentant plus de deux classes (Tableau 18). Ce fait est accentué lorsque le nombre de classes augmente. Pour un nombre de classes égal à deux, notre architecture ne se distingue pas de la structure d'un RMC classique, ce qui explique des résultats moindres par rapport à un RMC utilisant des techniques d'amélioration (algorithmes génétiques et sélection de caractéristiques). La sélection de caractéristiques par *SFFS* en *wrapper* ne fait qu'accroître le gain au niveau du taux de reconnaissance par rapport aux autres approches neuronales.

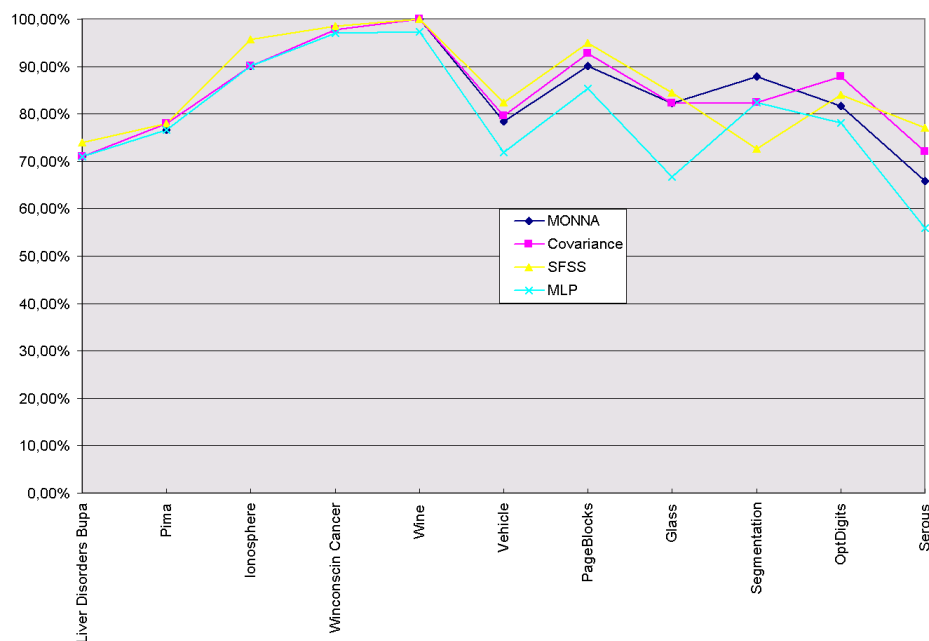


Figure 18. Comparaison des taux de reconnaissance de l'architecture MONNA avec et sans sélection d'attributs ainsi qu'avec un RMC.

Base	MONNA avec Sélection par <i>SFFS</i>	MONNA	ADHOC	DistAI	GADistAI
Pima	77,92%	76,6%	73,2%	76,3%	79,5%
Ionosphere	95,77%	90,1%	-	94,3%	98,6%
Winconsin Cancer	98,54%	97,8%	-	97,8%	99,3%
Wine	100%	100,0%	-	97,1%	99,4%
Vehicle	82,45%	78,4%	69,6%	65,4%	68,8%
Glass	84,44%	82,2%	70,5%	70,5%	80,8%

Tableau 18. Comparaison des taux de reconnaissance entre l'approche MONNA et d'autres approches neuronales.

On peut conclure que les méthodes de sélection d'attributs que nous avons mises en œuvre sont moins efficaces que celles utilisant les algorithmes génétiques. Une méthode de sélection d'attributs par un algorithme génétique semble donc ouvrir une voie prometteuse et serait une façon d'améliorer notre architecture.

4. Références

- [Almuallim 91] Almuallim H., Dietterich T.G., « Learning with many irrelevant features », *Ninth National Conference on Artificial Intelligence*, MIT press, pp. 547-552, 1991.
- [Almuallim 94] Almuallim H., Dietterich T.G., « Learning Boolean concepts in the presence of many irrelevant features », *Artificial Intelligence*, Vol. 69, n°1-2, pp. 279-305, 1994
- [Ash 89] Ash T., « Dynamic node creation in backpropagation networks », *Connection Science*, Vol. 1, n°4, pp. 365-375, 1989.
- [Blum 97] Blum A.L, Langley P., « Selection of relevant features and examples in machine learning », *Artificial Intelligence Journal special issue on relevance*, pp. 245-271, 1997.
- [Comon 91] Comon P., « Classification non supervisée par des réseaux multicouches », *Traitement du Signal*, Vol. 6, n°6, pp. 387-407, 1991.
- [Dash 97] Dash M., Liu H., « Feature selection for classification », *Intelligent Data Analysis*, Vol. 1, n°3, 1997.
- [Diday 82] Diday E., Lemaire J., Pouget J., Testu F., « *Eléments d'analyse de données* », Dunod, 1982.
- [Doak 92] Doak J., « An empirical evaluation of feature selection methods and their application to computer security », *Technical Report CSE-92-18*, Davis, University of California, Department of Computer Science, 1992.
- [Fahlman 89] Fahlman S.E., Lebiere C., « *The cascade-correlation learning architecture* », *Advances in Neural Information Processing Systems I*, D. Touretzky Editions, Morgan Kaufmann, San Mateo, CA, pp. 524-532, 1989.
- [Frean 90] Frean M., « The Upstart algorithm : a method for constructing and training feedforward neural networks », *Neural Computation*, Vol. 2, n° 2, pp. 198-209, 1990.
- [Fukunaga 90] Fukunaga K., « *Statistical Pattern Recognition* », Academic Press, 2nd Edition 1990.
- [Gennari 89] Gennari J.H., Langley P., Fisher D., « Models of incremental concept information », *Artificial Intelligence*, Vol. 40, pp. 11-61, 1989.
- [Hallinan 98] Hallinan J., Jackway P., « Cooperative evolution of a neural classifier and feature selection », *Proceedings of the second Asia pacific conference on simulated evolution and learning*, Vol. 2, pp. 137-143, 1998.
- [Herrault 94] Herrault J., Jutten C., « *Réseaux neuronaux et traitement du signal* », *Traité des nouvelles technologies, Série traitement du signal*, Ed. Hermès, 1994.

[John 94] John G.H., Kohavi R., Pfeger K., « *Irrelevant Features and the Subset Selection Problem* », in W. Cohen and H. Hirsh, editors, *Machine Learning : Proceedings of the eleventh International Conference*, Morgan Kaufmann Publishers, San Francisco, pp. 121-129, 1994.

[Jurie 99] Jurie F., Tomczak R., « Automatic determination of discriminant parameters. Application to the classification of Pollen », *QCAV'99*, Trois-Rivières, Canada, 1999.

[Kaynak 95] Kaynak C., « *Methods of Combining Multiple Classifiers and Their Applications to handwritten Digit Recognition* », M.Sc. Thesis, Institute of Graduate Studies in Science and Engineering, Bogazici University, 1995.

[Kira 92] Kira K., Rendell L., « A practical approach to feature selection », *Proceedings of the 9th International Conference on Machine Learning*, Morgan Kaufmann, pp. 249-256, 1996.

[Kittler 86] Kittler J., « *Feature selection and extraction* », Chapter 3, *Handbook of Pattern recognition and image processing*, Eds. Tzay Y. Young, King-Sun Fu, Academic Press, pp. 59-83, 1986.

[Kohavi 98] Kohavi R., John G.H., « *The wrapper approach* », H. Liu et H. Motoda, editors, *Feature Selection for Knowledge Discovery in Databases*, Springer Verlag, 1998.

[Kohavi 97] Kohavi R., John G.H., « Wrappers for feature selection », *Artificial Intelligence Journal special issue on relevance*, pp. 237-324, 1997.

[Kohavi 95] Kohavi R., Sommerfield D., « Feature subset selection using the wrapper method : overfitting and dynamic search space topology », *First International Conference on Knowledge Discovery and Data Mining*, 1995.

[Kohavi 94] Kohavi R., « Feature subset selection as search with probabilistic estimates », *AAAI Fall Symposium on Relevance*, pp. 122-126, 1994.

[Langley 94] Langley P., « Selection of relevant features in machine learning », *In Proceedings of the AAAI Fall Symposium on Relevance*, pp.1-5, 1994.

[Lezoray 98] Lezoray O., Elmoataz A., Cardot H., Gougeon G., Lecluse M., Elie H., Revenu M., « Segmentation d'images couleur application à la cytologie en vue d'un tri informatique cellulaire », *Actes de RFIA '98*, Clermont-ferrand, France, Janvier 1998.

[Lezoray 99a] Lezoray O., Elmoataz A., Cardot H., Gougeon G., Lecluse M., Elie H., Revenu M., « Segmentation of cytological images using color and mathematical morphology », *Acta Stereologica*, Vol. 18, n°1, 1999.

[Lezoray 99b] Lezoray O., Elmoataz A., Cardot H., Revenu M., « ARCTIC : An automatic system for cellular sorting by image analysis », *Proceedings of Vision Interface 99*, Trois-Rivières, Canada, 1999.

- [Liu 96a] Liu H., Setiono R., « Feature selection and classification - a probabilistic wrapper approach », *Proceedings of the Ninth International Conference on Industrial and Engineering Applications of AI and ES*, 1996
- [Liu 96b] Liu H., Setiono R., « A probabilistic approach to feature selection - a filter solution », *Proceedings of the Thirteenth International conference on Machine Learning*, Morgan Kaufmann, 1996.
- [Malerba 95] Malerba D., Esposito F., Semeraro G., « A Further Comparison of simplification Methods for Decision-Tree Induction », in D. Fisher and H. Lenz (Eds.), *Learning from Data: Artificial Intelligence and Statistics V*, Lecture Notes in Statistics, Springer Verlag, Berlin, 1995.
- [Mangasarian 90] Mangasarian O.L., Wolberg W.H., « Cancer diagnosis via linear programming », *SIAM news*, Vol. 23, n°5, pp.1-18, 1990.
- [Mezard 89] Mezard M., Nadal J.P., « Learning in feedforward layered networks : the tiling algorithm », *Journal of Physics A*, Vol. 2, n° 12, pp. 2191-2203, 1989.
- [Murphy 94] Murphy P.M., Aha D.W., UCI Repository of machine learning databases : Machine-readable data repository, Irvine, CA : University of California, <<http://www.ics.uci.edu/~mlearn/MLRepository.html>>, Department of Informatics and Computer Science, 1994.
- [Narendra 77] Narendra P.M., Fukunaga K., « A branch and bound algorithm for feature subset selection », *IEEE transactions on Computers*, Vol. 26, n°9, pp. 917-922, 1977.
- [Pudil 94] Pudil P., Novovocova J., Kittler J., « Floating search methods in feature selection », *Pattern recognition letters*, Vol. 15, pp. 1119-1125, 1994.
- [Quinlan 93] Quinlan R., « *C4.5 : Programs for Machine Learning* », Morgan Kaufmann, 1993.
- [Richeldi 96] Richeldi M., Lanzi P., « Performing effective feature selection by investigating the deep structure of the data », *Proceedings of the second International Conference on Knowledge Discovery and Data Mining*, pp. 379-383, 1996.
- [Rissanen 1978] Rissanen J., « Modeling by shortest data description », *Automatica*, Vol. 14, pp. 465-471, 1978.
- [Sheinvald 90] Sheinvald J., Dom B., Niblack W., « A modeling approach to feature selection », *In Proceedings of the Tenth International Conference on Pattern Recognition*, pp. 535-539, 1990.
- [Siedlecki 88] Siedlecki W., Slansky J., « On automatic feature selection », *International Journal of Pattern recognition*, Vol. 2 , pp. 197-220.
- [Siedlecki 89] Siedlecki W., Slansky J., « A note on genetic algorithms for large-scale feature selection », *IEEE transactions on computers*, Vol. 10, pp. 335-347, 1989.

[Sigillito 89] Sigillito V.G., Wing S.P., Hutton L.V., Baker K.B., « Classification of radar returns from the ionosphere using neural networks », *Johns Hopkins APL Technical Digest*, Vol. 10, pp. 262-266, 1989.

[Smith 88] Smith J.W., Everhart J.E., Dickson W.C., Knowler W.C., Johannes R.S., « Using the ADAP learning algorithm to forecast the onset of diabetes mellitus », *Proceedings of the Symposium on Computer Applications and Medical Care*, IEEE Computer Society Press, pp. 261-265, 1988.

[Tenorio 90] Tenorio M.F., Lee W., « Selforganizing network for optimum supervised learning », *IEEE Transactions on Neural Networks*, Vol. 1, n° 1, pp. 100-110, 1990.

[Tomczak 99] Tomczak R., Jurie F., Rouquet R., Bonton P., « Automated measurements of Pollen Rates by computer Vision », *QCAV'99*, Trois-Rivières, Canada, 1999.

[Vafaie 93] Vafaie H., De Jong K., « Robust feature selection algorithms », *Proceedings of the 5th International Conference on Tools with artificial intelligence*, 1993.

[Yacoub 97] Yacoub M., Bennani Y., « HVS : a heuristic for variable selection in multilayer artificial neural network classifier », *Proceedings of ANNIE'97*, pp. 527-532, 1997.

[Yang 98a] Yang J., Parekh R., Honavar V., « DistAI : An inter-pattern distance based constructive learning algorithm », *Proceedings of the International Joint Conference on Neural Networks*, Anchorage, Alaska, 1998.

[Yang 98b] Yang J., Honavar V., « Feature subset selection using a genetic algorithm », *IEEE Intelligent Systems*, Vol. 13, pp. 44-49, 1998.

Application en microscopie

« Tout homme tend à devenir machine, habitude, méthode, maîtrise, enfin cela veut dire machine... En vérité c'est la machine qui tend à devenir animal ... »

Paul Valéry, Cahier, La Pléiade, pp. 885, 1975.

1. UN SYSTÈME DE TRI CELLULAIRE	158
1.1. GÉNÉRALITÉS D'ANATOMIE ET DE CYTOLOGIE PATHOLOGIQUES	158
1.1.1. L'ANATOMIE PATHOLOGIQUE	158
1.1.2. HISTORIQUE DE LA CYTOLOGIE	158
1.1.3. STRUCTURE ANATOMIQUE ET CYTOLOGIQUE	159
1.1.4. LES PRÉLÈVEMENTS EN ANATOMIE ET CYTOLOGIE PATHOLOGIQUES	161
1.1.5. LES COLORATIONS	162
1.1.6. TYPES CELLULAIRES EN CYTOLOGIE DES SÉREUSES	162
1.2. LE SCREENING	165
1.2.1. LE SCREENING MANUEL	165
1.2.2. LE SCREENING AUTOMATIQUE	166
1.3. LE SYSTÈME ARCTIC	168
2. L'ACQUISITION DES IMAGES	170
2.1. RÉSULTATS EXPÉRIMENTAUX SUR LE SYSTÈME	170
2.1.1. INFLUENCE DU TEMPS DE RÉPONSE CAMÉRA PLUS MICROSCOPE	170
2.1.2. INFLUENCE DE L'ENVIRONNEMENT	172
2.2. DÉTERMINATION DU TYPE DE BRUIT DU SYSTÈME	173
2.2.1. RAPPELS GÉNÉRAUX SUR LE BRUIT	173
2.2.2. BRUIT D'ACQUISITION DE NOTRE SYSTÈME	174
2.3. ÉLIMINATION DU BRUIT	174
3. RÉSULTATS EXPÉRIMENTAUX	176
3.1. ÉVALUATION DE LA SEGMENTATION	176
3.2. CONSTRUCTION D'UNE BASE DE DONNÉES	178
3.2.1. L'ÉTIQUETAGE PAR UN EXPERT	178
3.2.2. INTERSECTION ENTRE PLUSIEURS EXPERTS	180
3.3. CLASSIFICATION DES OBJETS	181
4. FONCTIONNALITÉS DU SYSTÈME ARCTIC	183
5. CONCLUSION	187

Figures et Tableaux

FIGURE 1. REPRÉSENTATION SCHÉMATIQUE DE LA CELLULE [TAKAHASHI 83].	159
FIGURE 2. EPITHÉLIUM PAVIMENTEUX STRATIFIÉ [MENARD 93].	160
FIGURE 3. DIFFÉRENTS TYPES D'ÉPITHÉLIUM SUIVANT LES ORGANES [MENARD 93].	160
FIGURE 4. SCHÉMA DES CAVITÉS SÉREUSES DU CORPS HUMAIN. LES CAVITÉS PLEURALES, PÉRICARDIQUE ET PÉRITONÉALE SONT COMPRISES ENTRE UN FEUILLET PARIÉTAL ET UN FEUILLET VISCÉRAL. 1 : PLÈVRE PARIÉTALE, 2 : PLÈVRE VISCÉRALE, 3 : PÉRICARDE, 4 : ÉPICARDE, 5 : DIAPHRAGME, 6 : PÉRITOINE PARIÉTAL, 7 : PÉRITOINE VISCÉRAL [TAKAHASHI 83].	161
FIGURE 5. QUELQUES TYPES CELLULAIRES EN CYTOLOGIE DES SÉREUSES. LES CELLULES PEUVENT ÊTRE ISOLÉES OU EN AMAS.	164
FIGURE 6. LE <i>SCREENING</i> MANUEL SUR UNE LAME.	165
FIGURE 7. UTILISATION D'ARCTIC POUR LE CONTRÔLE DE QUALITÉ.	169
FIGURE 8. LE SYSTÈME D'ACQUISITION D'IMAGES MICROSCOPIQUES.	170
FIGURE 9. MOYENNE DU BRUIT SUR LES COMPOSANTES R, V ET B EN FONCTION DE L'HEURE D'ACQUISITION.	172
FIGURE 10. VARIANCE DU BRUIT SUR LES COMPOSANTES R, V ET B EN FONCTION DE L'HEURE D'ACQUISITION.	172
FIGURE 11. ERREUR MOYENNE SUR CHAQUE COMPOSANTE EN FONCTION DES SOURCES LUMINEUSES UTILISÉES.	173
FIGURE 12. NOMBRE DE PIXELS EN FONCTION DU BRUIT SUR LA COMPOSANTE VERTE.	174
FIGURE 13. UNE IMAGE DES SÉREUSES AVANT (A) ET APRÈS (B) DIVISION PAR UNE IMAGE DE FOND.	175
FIGURE 14. DES ERREURS DE SEGMENTATION.	178
FIGURE 15. LA CONSTITUTION DE BASES D'OBJETS PAR UN EXPERT.	179
FIGURE 16. LE LOGICIEL ARCTIC.	184
FIGURE 17. LE PROCESSUS D'ÉVALUATION DE LA SEGMENTATION D'UNE IMAGE.	184
FIGURE 18. LA CONSULTATION DES CARACTÉRISTIQUES DES CELLULES D'UNE IMAGE.	185
FIGURE 19. LES BOITES DE DIALOGUE D'ÉTIQUETAGE D'UNE IMAGE DE CELLULES.	185
FIGURE 20. LA CONSULTATION D'UNE BASE DE CELLULES.	186
TABLEAU 1. TAUX DE SEGMENTATION CORRECTE POUR LES CELLULES.	176
TABLEAU 2. TAUX D'ERREUR DE LA SEGMENTATION SELON LE TYPE D'ERREUR.	177
TABLEAU 3. LES CLASSES DE CELLULES.	179
TABLEAU 4. LES TAUX DE CONCORDANCE ENTRE LES EXPERTS SUIVANT LA BASE UTILISÉE.	180
TABLEAU 5. LA MATRICE DE CONFUSION.	182
TABLEAU 6. LA MATRICE DE CONFUSION DU TABLEAU 5 APRÈS REGROUPEMENT.	182

Dans cette troisième et dernière partie, nous présentons le problème du tri cellulaire en anatomie et cytologie pathologiques pour la cytologie des séreuses. Après quelques rappels de cytologie, nous présentons le problème à résoudre, à savoir la construction d'un système de détection de cellules anormales. Nous détaillons ensuite le fonctionnement d'un tel système et nous présentons les conditions d'acquisition des images. Enfin nous présentons les résultats du système ainsi que ses fonctionnalités logicielles.

1. Un système de tri cellulaire

1.1. Généralités d'anatomie et de cytologie pathologiques

1.1.1. L'anatomie pathologique

L'anatomie pathologique est l'application aux cellules et aux tissus prélevés chez l'homme de diverses méthodes d'analyse basées principalement sur la morphologie, à des fins de diagnostic, de pronostic et de meilleure compréhension des causes et mécanismes des maladies. La pathologie est l'étude et la science des maladies. La traduction littérale de pathologie en grec (pathos, logos) est « les mots de la souffrance ». Le pathologiste étudie les causes et les conséquences de la maladie, comment la maladie affecte certains tissus ou cellules, la progression de la maladie dans le corps humain, comment la maladie se manifeste et enfin les méthodes pour surveiller la progression de la maladie. La pathologie donne un fondement scientifique pour la médecine clinique et sert de lien entre les sciences fondamentales et les soins au patient.

L'étude morphologique des lésions qui sont la cause ou la conséquence des maladies peut se faire :

- à l'œil nu : c'est l'anatomie pathologique macroscopique,
- au microscope :
 - c'est l'histologie pathologique (histo signifie tissu en grec) où l'on examine des coupes de tissus colorées,
 - c'est la cytologie pathologique (cyto signifie cellule en grec) où l'on examine un étalement de cellules isolées.

Le domaine que nous allons étudier, la cytologie, fait partie intégrante de la pathologie. La cytologie a connu un développement considérable grâce à des techniques d'observation de plus en plus sophistiquées. Ces techniques permettent d'obtenir une multitude d'images (microscopie photonique, à contraste de phase, électronique).

1.1.2. Historique de la cytologie

En 1838, Mueller [Mueller 1838] fut le premier à décrire et illustrer des cellules malignes de tumeurs observées au microscope. Les premiers atlas de cytologie furent réalisés par des microscopistes français, Donné [Donné 1845] et Lebert [Lebert 1845] ainsi qu'un atlas de cytologie sur le cycle menstruel réalisé par Pouchet [Pouchet 1847]. Des rapports sur les tumeurs en cytologie ont été réalisés ensuite par Beale [Beale 1860] sur des crachats, par Sanders [Sanders 1864] sur de l'urine, par Dufour [Dufour 1860] sur des fluides cérébraux, par Marissi [Marissi 1909] sur des spécimens gastriques. En 1928, Papanicolaou [Papanicolaou 28] présenta ses travaux sur l'étude cytologique du tractus génital féminin ouvrant ainsi la voie vers le dépistage systématique du cancer dans les années 1940.

1.1.3. Structure anatomique et cytologique

On peut distinguer quatre niveaux structuraux : les cellules, les tissus, les organes et les appareils ou systèmes.

A son niveau structural le plus simple, l'organisme est fait de cellules. La cellule (Figure 1) est l'unité fonctionnelle la plus petite de l'organisme. Les cellules peuvent être différentes entre elles selon leur fonction.

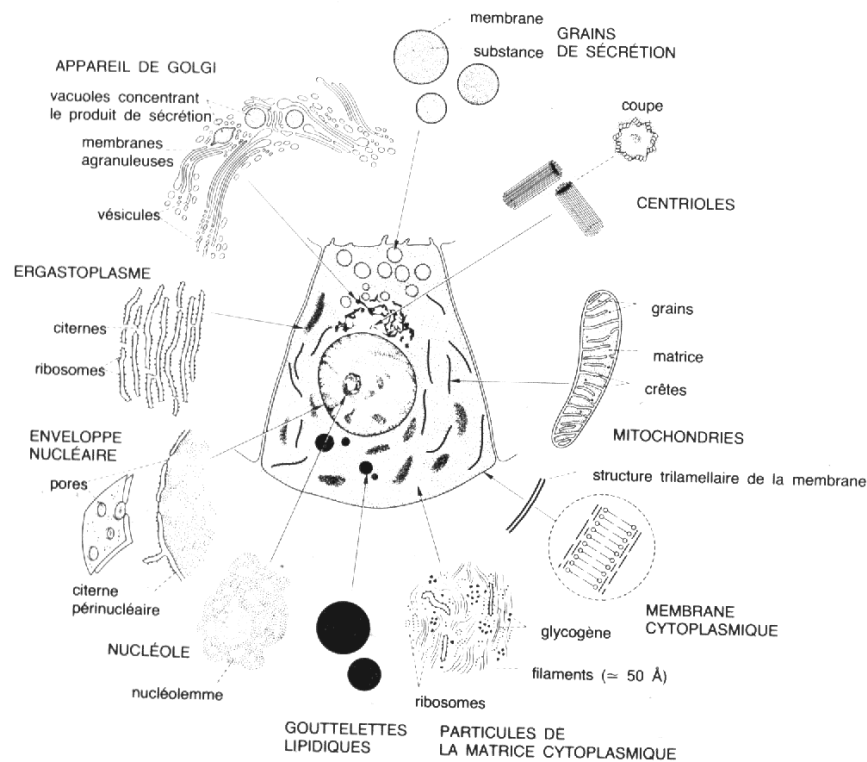


Figure 1. Représentation schématique de la cellule [Takahashi 83].

Chez la majorité des animaux, y compris chez l'homme, des groupes de cellules semblables s'associent pour former des tissus. On peut définir un tissu comme un groupement de cellules qui possèdent fondamentalement la même structure et qui assurent un même ensemble de fonctions simples. On distingue ainsi [ENSTAC 87] :

1. les tissus épithéliaux : ils recouvrent la surface des organes,
2. les tissus conjonctifs : ils servent aussi bien de soutien (os, cartilages) que d'attache à d'autres tissus (tendons, ligaments ...),
3. les tissus musculaires composés de cellules spécialisées capables de se contracter,
4. les tissus nerveux : l'encéphale, la moelle épinière et les nerfs par exemple sont formés de tissus nerveux. Ces cellules sont caractérisées par de longs prolongements,
5. le tissu hématopoïétique (ou tissu sanguin).

Nous nous limiterons à l'étude des tissus épithéliaux. Ils sont caractérisés par une étroite juxtaposition des éléments cellulaires et ils reposent toujours sur une membrane basale : couche de collagène qui sépare le chorion de l'épithélium (Figure 2).

Il existe différents types d'épithélium se différenciant par leur structure :

- l'**épithélium pavimenteux simple** possède une seule couche de cellules aplaties. Le mésothélium en est un.
- l'**épithélium pavimenteux stratifié** possède plusieurs couches dont le nombre varie d'une région à l'autre. Ce type de tissu recouvre les parois de la bouche, du pharynx.
- l'**épithélium de transition** est un tissu stratifié spécialisé tapissant.
- l'**épithélium cubique simple** compose les glandes ou bien la thyroïde.
- l'**épithélium cylindrique simple** tapisse le tube digestif ou bien l'estomac.



Figure 2. Epithélium pavimenteux stratifié [Menard 93].

La figure 3 représente les différents types d'épithélium suivant les organes.

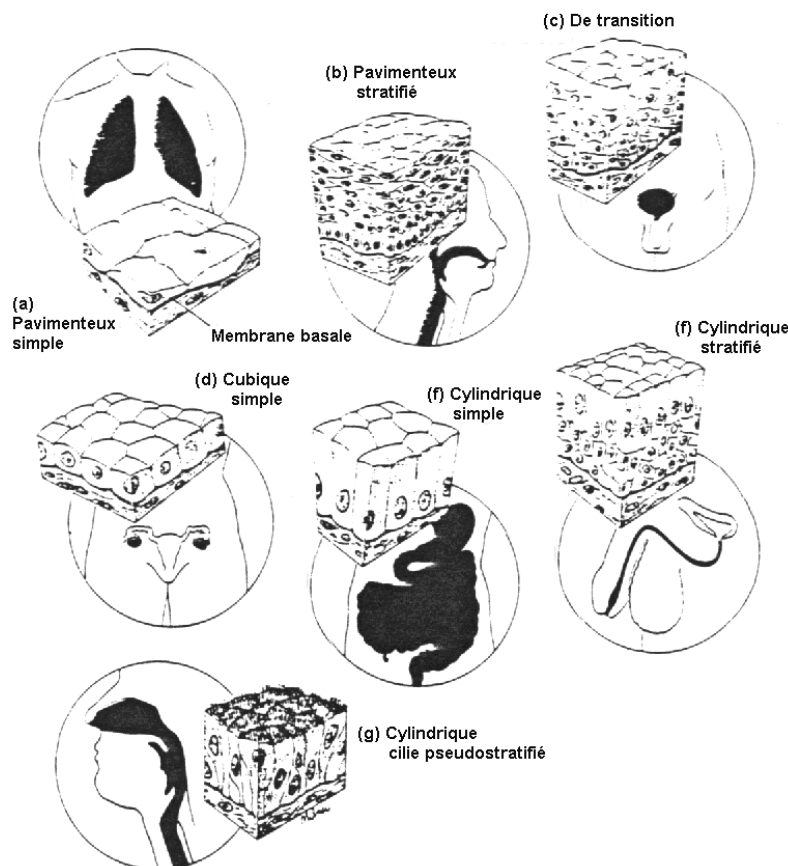


Figure 3. Différents types d'épithélium suivant les organes [Menard 93].

Nous pouvons définir les organes comme l'assemblage de plusieurs tissus permettant d'accomplir des fonctions plus complexes. Par exemple, l'estomac est tapissé de tissus épithéliaux, il est entouré de tissus musculaires et innervé par du tissu nerveux. Des tissus conjonctifs les maintiennent ensemble.

On entend enfin par appareil l'ensemble des organes qui accomplissent une tâche commune.

1.1.4. Les prélèvements en Anatomie et Cytologie Pathologiques

Un prélèvement nécessite un personnel qualifié et un matériel adapté. C'est un acte médical et il ne doit pas être traumatisant pour les cellules. On peut distinguer plusieurs types de prélèvements :

- **les biopsies** : c'est un examen pratiqué sur un être vivant. En fait, on réserve ce terme aux petits prélèvements faits dans le but d'obtenir un diagnostic,
- **les exérèses** : ce terme concerne les prélèvements de pièces plus importantes quand on a cherché à retirer la totalité de la lésion,
- **les autopsies** : elles servent à déterminer les causes de la mort et les pathologies associées.
- Les prélèvements cytologiques :
 - les liquides d'épanchement des séreuses (Figure 4),
 - les urines,
 - les liquides céphalo-rachidiens,
 - les expectorations, aspirations bronchiques, brossages,
 - les frottis génitaux,
 - les cytoponctions.

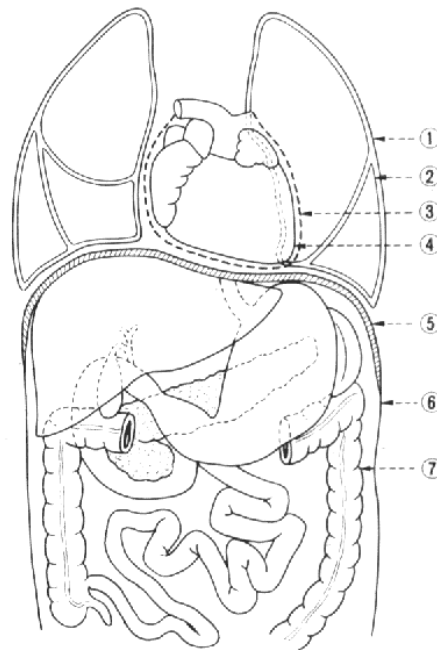


Figure 4. Schéma des cavités séreuses du corps humain. Les cavités pleurales, péricardique et péritonéale sont comprises entre un feuillet pariétal et un feuillet viscéral. 1 : plèvre pariétale, 2 : plèvre viscérale, 3 : péricarde, 4 : épicarde, 5 : diaphragme, 6 : péritoine pariétal, 7 : péritoine viscéral [Takahashi 83].

Après le prélèvement on observe une autolyse des cellules (dégénérescence) et pour les conserver, on procède à une fixation. Les fixateurs sont de deux sortes : les fixateurs chimiques (Alcool-Ether, Spray, etc.) ou physiques (séchage à l'air). Nous travaillons sur les liquides d'épanchement des séreuses.

L'échantillon est « techniqué » par centrifugation afin d'obtenir une concentration cellulaire. Le culot de centrifugation est étalé sur des lames puis fixé et coloré. La fixation

utilisée dans le cadre de notre étude reste dans le domaine d'une fixation standard (Alcool-Acétone).

1.1.5. Les colorations

Il n'existe pas de technique de coloration polyvalente. Selon le but à atteindre, en l'occurrence selon la nature chimique des différents constituants cellulaires à mettre en évidence, on fera appel à des techniques différentes.

Dans un grand nombre de cas les méthodes de coloration de topographie générale sont suffisantes. Parfois cependant des techniques histochimiques sont indispensables pour préciser le diagnostic. La coloration ne peut être réussie que si elle est précédée d'une fixation appropriée visant à préserver l'intégrité des structures à colorer.

En cytologie les fixateurs essentiellement utilisés sont le mélange alcool-éther ou des laques fixantes, ces deux méthodes permettant ensuite une coloration standard. D'autres étalements sont séchés sous réserve d'effectuer des colorations particulières telles que le May-Grunwald-Giemsa, le Bleu Alcian, le Perls et le P.A.S.

Les principales colorations cytologiques sont :

- la coloration de Papanicolaou [Papanicolaou 47,54],
- la coloration de Harris-Shorr,
- le May-Grunwald-Giemsa,
- la coloration de Perls,
- le P.A.S,
- le Bleu-Alcian.

Le standard international est la coloration de Papanicolaou qui permet d'obtenir les colorations suivantes. Les noyaux cellulaires sont colorés en bleu violet par l'hématoxyline de Harris, les cytoplasmes en vert ou bleu vert par EA50 et OG6. Nous avons utilisé ce standard pour colorer les cellules. Les colorations en cytologie sont nombreuses, nous avons utilisé le standard international de coloration car les pathologistes possèdent une certaine expérience dans la reconnaissance de cellules des séreuses colorés de cette manière et un changement de coloration aurait entraîné un apprentissage difficile.

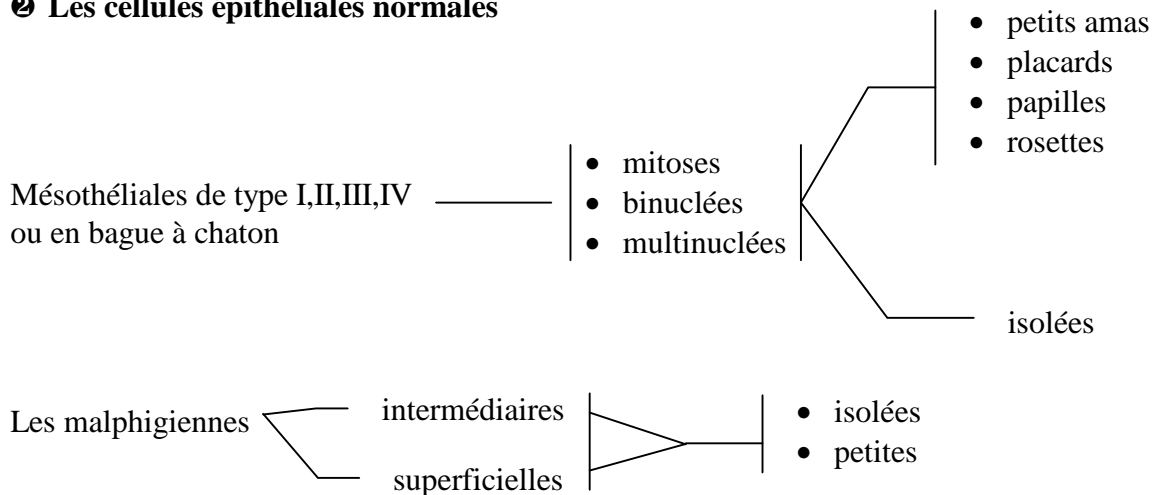
1.1.6. Types cellulaires en cytologie des séreuses

Les différents types de prélèvements possibles ainsi que les différentes colorations ont été présentés. Nous travaillons sur des cellules provenant de la cytologie des séreuses colorées par le standard international de Papanicolaou. Afin de bien connaître tous les objets cellulaires pouvant être rencontrés, ceux-ci ont été répertoriés. Ils se répartissent en 6 catégories.

❶ Les objets du fond de la préparation

- les hématies ou globules rouges,
- les débris cytoplasmiques,
- les débris nucléaires,
- les noyaux nus,
- les amas de mucus, les spires de Curshmann ou les calcosphérites,
- les cristaux de Charcot-Leyden extra-cellulaires,
- les éléments mycéliens,
- les substances étrangères (grains de colorants, poussières, cellules végétales, grains de talc, dépôts de carbone des mines de crayon, bulles d'air, alternaria).

② Les cellules épithéliales normales



③ Les cellules inflammatoires

Les polynucléaires

- neutrophiles
- éosinophiles
- altérés (pyocytes)

Les lymphocytes

- activés
- non activés

Les histiocytes

Les macrophages

Sidérophages

Anthracosés

④ Les cellules tumorales primaires

Cellules du mésothéliome malin

- isolées
- en papilles

⑤ Les cellules tumorales secondaires

Cellules adénocarcinomateuses
(carcinomes glandulaires)

- isolées
- isolées en bague à chaton
- en morules
- en papilles

- peu de mucus
- beaucoup de mucus

Cellules des carcinomes épidermoïdes
(carcinomes malphigiens)

- isolées
- en petits amas

Plus ou moins
kératinisées

⑥ Les cellules tumorales rares

Cellules des lymphomes malins en cellules isolées,
Sarcomes en cellules isolées,

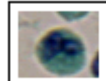
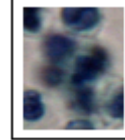
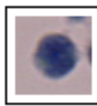
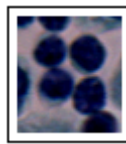
Les mélanomes malins

- cellules isolées
- en petits amas

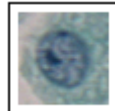
Voici quelques exemples de types cellulaires (isolés ou en amas) en cytologie des séreuses (Figure 5). On remarque la diversité des configurations tant au niveau de la couleur (de pâle à foncé) que de la répartition spatiale pour des cellules en amas.



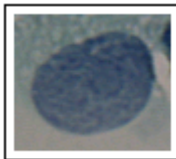
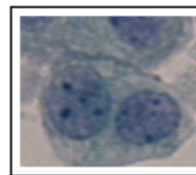
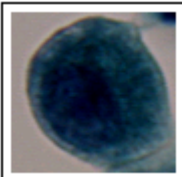
hématie

Polynucléaire
neutrophilePolynucléaire
éosinophilePolynucléaire
altéréAmas de
polynucléaires
neutrophilesLymphocyte
activéLymphocyte non
activéAmas de lymphocytes
non activés

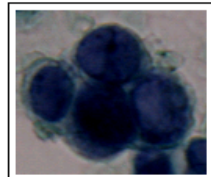
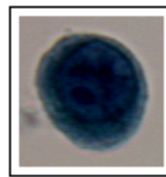
Macrophage



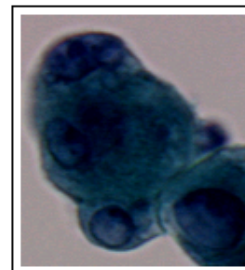
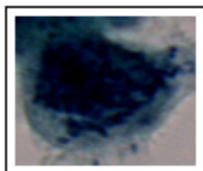
Mésothéliale

Mésothéliale
dystrophiqueAmas de
mésothéliales

Adénocarcinome

Amas de cellules
adénocarcinomateuses

Mésothéliome

Amas de cellules
du mésothéliome

Mitose anormale

Figure 5. Quelques types cellulaires en cytologie des séreuses. Les cellules peuvent être isolées ou en amas.

1.2. Le screening

1.2.1. Le screening manuel

Une fois le prélèvement concentré, étalé sur une lame, les cellules sont colorées par la coloration de Papanicolaou [Papanicolaou 47,54]. La phase d'expertise de la lame peut commencer. Cette phase s'appelle le *screening*. Cela consiste en une étape de lecture de la lame par un expert en cytologie (un cytotechnicien). Il repère les cellules intéressantes, à savoir les cellules qui sont ou qui pourraient être anormales. Le *screening* est donc un repérage et une évaluation visuelle de toutes les cellules présentes sur une lame afin de détecter et d'identifier d'éventuelles anomalies. Son but est principalement la détection des cellules anormales ou suspectes afin d'établir un diagnostic. Le diagnostic le plus important étant la présence de cellules anormales. Ceci est donc d'un intérêt capital pour le pathologiste dont le diagnostic dépend de la bonne détection et reconnaissance des cellules anormales ou suspectes lors du *screening*.

Un *screening* s'effectue de la manière suivante : le cytotechnicien parcourt la lame de gauche à droite et de haut en bas (technique d'aller-retour). A chaque instant il inspecte une partie de la lame à travers un champ microscopique (Figure 6). Lorsque le cytotechnicien passe au champ suivant, il procède à une superposition entre les champs (précédent et actuel). Cette superposition est horizontale et verticale (les superpositions sont données par les zones hachurées) . Ceci permet de sécuriser le *screening* et de rendre le dépistage plus fiable.

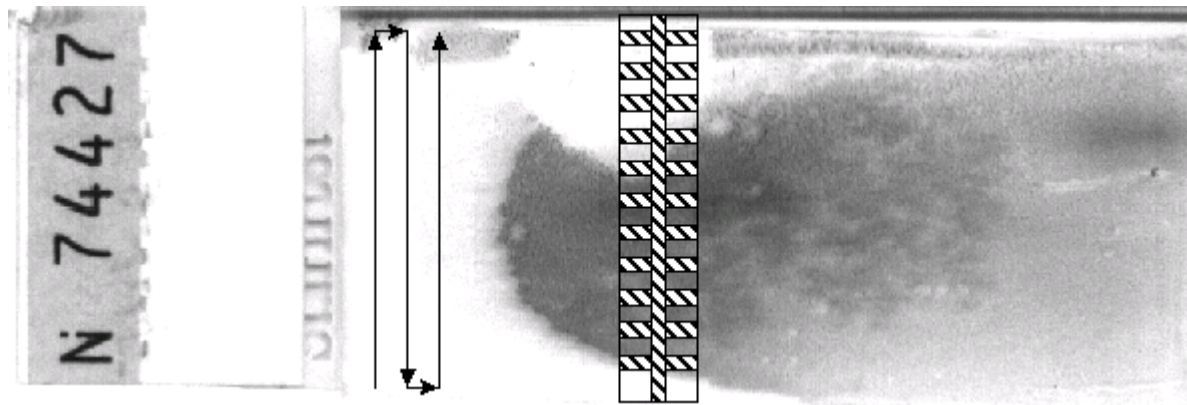


Figure 6. Le screening manuel sur une lame.

Le *screening* manuel des prélèvements cytologiques est donc difficile, fatigant, consommateur de temps et est décrit comme un travail « intense, particulièrement complexe dont le résultat repose sur le point de vue de l'être humain » [Koss 89]. Le faible nombre de cellules anormales (comparativement aux cellules normales) observées durant le *screening* implique une grande concentration du cytotechnicien. Un cytotechnicien voit entre 30 et 200 lames par jour. Sa culture médicale est remise en cause chaque jour. A cause de ce facteur subjectif inévitable, quelques erreurs peuvent apparaître et causer des faux négatifs (une lame donnée comme ne contenant pas de cellules anormales alors que c'est le cas). Or les préoccupations principales en cytologie sont celles qui permettent d'augmenter le potentiel de vérité contenu dans le diagnostic. Un cytopathologiste s'attache donc à réaliser le moins de faux positifs possible (une lame donnée comme anormale alors qu'elle ne contient que des cellules normales) mais surtout de faux négatifs. Des taux de faux négatifs ont été donnés

dans la littérature et certains travaux alarmants [Koss 93] font état de 50% de faux négatifs pour la cytologie du col. Les principales causes d'erreur en cytologie sont parmi les suivantes :

- Les prélèvements non significatifs (mal faits, mal fixés, acellulaires),
- Les erreurs de mauvaise interprétation
- Les erreurs d'omission par diminution de la vigilance,
- Le grand nombre de lames à analyser.

Certaines mesures ont été envisagées pour effectuer une vérification de 10% des lames données comme négatives : cela a été prescrit par le gouvernement américain par arrêté fédéral en réponse à un contrôle de qualité dans les services de cytopathologie [Revision 67]. Cependant après plusieurs années, il n'est pas apparu que ce genre de contrôle des lames élimine des erreurs de *screening* [Koss 93]. La relecture rapide de toutes les lames donne de meilleurs résultats mais reste consommatrice de temps.

1.2.2. Le screening automatique

Afin d'améliorer le contrôle de la qualité du *screening*, certaines recherches se sont orientées vers une automatisation totale du *screening*. La mise en place d'un système informatique peut apporter au cytopathologiste un plus dans le domaine de l'assurance de la qualité. En effet l'examen des lames par un cytotechnicien est une tâche subjective et celle-ci n'est pas infaillible. Un système informatique peut permettre de donner une analyse objective de la lame. De tels systèmes basés sur l'analyse d'images sont apparus rapidement.

1.2.2.1. Historique des systèmes de screening

Historiquement, la première tentative d'automatisation par analyse d'images a été tentée par Tolles avec le cytoanalyser [Tolles 55] mais ce fut un échec. Quelques années plus tard apparut le CYDAC (Cytometric Data Conversion) de Prewitt J.M.S et Mendelsohn [Prewitt 66] pour la classification de leucocytes. TICAS fut proposé en 1968 par Wied G.L [Wied 68] pour l'analyse des cellules dans des frottis vaginaux. On remarquera que les travaux de recherche portent principalement sur les frottis car c'est ce type d'examen qui fournit le plus de lames en cytologie (environ 80%). Ces débuts d'informatisation et d'automatisation par analyse d'image pour les cellules de frottis vaginaux ont été suivis par des essais pour tenter d'automatiser complètement le *screening* des lames de frottis. L'idée était d'automatiser un travail assez monotone par une machine qui non seulement serait plus fiable que le cytopathologiste mais qui donnerait également des résultats valables en un temps rapide et de façon économique. Un des seuls systèmes ayant pris en compte réellement ce genre de contraintes est le système CYBEST (Japon) [Tanaka 77,87] dont la construction a commencé en 1967 et qui a continué jusqu'en 1986. Ce système regroupait toutes les technologies nécessaires à ce genre d'investigations : micro-ordinateur, scanner vidéo, observation optique, auto changeur de lames. Le premier prototype apparut en 1972 et était dédié à la détection du cancer du col de l'utérus. Bien que prometteur, le fait d'avoir voulu tout intégrer dans la même machine la rendait trop lente à cause de la préparation de la lame. Mais c'est surtout l'absence de coopération entre le système et l'opérateur humain pour la donnée du diagnostic qui est la raison majeure d'échec des systèmes entièrement automatiques. Car dans ces systèmes la séparation entre spécimens normaux et anormaux n'était pas suffisamment efficace et reproductible. La plupart des cellules détectées comme anormales étaient dues à des artefacts. Parmi les systèmes informatiques automatiques qui ont été développés, on peut également citer BioPEPR (Pays Bas) [Zanisher 82], CERVIFIP (Royaume Uni) [Tucker 82], DIASCANNER (Suède) [Nordin 89], FAZYTAN (Allemagne) [Erhardt 80] et LEYTAS (Pays Bas) [Ploem 87].

C'est pourquoi un système entièrement automatique n'a encore jamais vu le jour et les cytopathologistes se sont plutôt tournés vers le développement d'un système semi-automatique qui laisserait le jugement à l'expert. Le rôle de ce type de système est alors de détecter les lames suspectes (à savoir celles présentant des cellules atypiques ou anormales) et de présenter ces cellules au pathologiste pour qu'il puisse juger la lame. Un tel système est un système de *rescreening*. Le système peut alors soit amener l'objectif du microscope à l'endroit où se trouvent les cellules "anormales" ou bien les afficher sur un écran en précisant les cellules à diagnostiquer. Le rôle de la machine ne serait plus alors de diagnostiquer une lame cytologique mais plutôt de présenter au cytopathologiste les zones suspectes pour un jugement. Ainsi, on ne cherche pas à remplacer le travail du cytopathologiste mais à le faciliter et à le sécuriser.

Nous allons maintenant faire un panorama exhaustif des différents travaux de recherche sur les systèmes semi-automatiques effectués jusqu'à nos jours :

- Le système meso-TICAS [Dytch 87] qui est la suite du TICAS développé par Wied G.L. Il fut développé en 1986 à l'Université de Chicago [Bartels 87],
- Le système SCIMA (System for Cytologic Image Analysis) [Hall 87] développé à l'université de Los Angeles de 1977 à 1985. Il est principalement adapté à la détection de lésions précancéreuses du col de l'utérus,
- Le système IRIS [Kasdan 87] réalisé par International Remote Imaging Systems en 1987. Son principal attrait réside dans la sélection des zones d'intérêt par le cytopathologiste grâce à un écran sensible,
- Le système NeoPath [Patten 96] est un système de contrôle de qualité qui vérifie toutes les lames classées normales et satisfaisantes par les cytologistes pour aider à éviter les faux négatifs.
- Le système ROCHE [Knesel 96] qui fonctionne sur des lames préparées par un procédé amélioré permet d'obtenir une préparation homogène avec une bonne coloration. Le système réalise un *screening* de la lame et présente les cellules suspectes ou anormales au cytotechnicien.
- Le système HOME [Derder 95] qui introduit un nouveau concept : les systèmes ergonomiques intégrés. Ce type de système permet de fournir des outils de haut niveau nécessaires aux tâches d'observation, de sélection, de marquage, etc. Il permet d'assister le cytopathologiste et de contrôler ses activités de diagnostic. Des événements représentent les activités d'interaction du cytopathologiste avec le système, ce qui permet de modéliser l'utilisateur à partir de son activité sous le système. L'objectif de cette modélisation étant d'adapter le comportement du système à l'évolution des caractéristiques de cet utilisateur,
- Le système PAPNET [Koss 93,94, Boon 93, Mango 93, Sherman 94, Rosenthal 96]. Il n'opère que sur des frottis et sa particularité réside dans l'utilisation de réseaux neuronaux pour la classification des cellules, ce qui a donné des résultats 20% meilleurs pour la plupart des frottis examinés.

1.2.2.2. Structure d'un système informatique de *rescreening*

Les systèmes informatique de *rescreening* sont conçus pour opérer de façon postérieure au *screening* manuel. Mais ils peuvent éventuellement être utilisés pour fonctionner de manière antérieure, cependant les pathologistes ont une préférence pour le *rescreening* informatique. L'architecture des systèmes utilisés pour le *rescreening* de lames de cytologie reste identique quel que soit le système employé et est constituée de trois étapes séquentielles:

- La segmentation : elle consiste à extraire des objets (les composantes cellulaires) à partir d'images couleur.
- La caractérisation : elle fournit des caractéristiques des objets précédemment extraits. Ces données doivent être suffisamment pertinentes et précises pour décrire les objets.
- La classification : les cellules sont classées dans un type cellulaire grâce à une méthode de classification (arbre de décision, réseaux de neurones, etc.). Cela consiste à déterminer le type de la cellule à classer et plus précisément à déterminer son caractère bénin ou malin.

1.3. Le système ARCTIC

Nous nous proposons de réaliser un système d'aide au *screening* manuel par le *rescreening informatique*. Ce système fonctionnera sur des lames provenant de la cytologie des séreuses (liquides d'épanchement de la plèvre et du péritoine). Travailler sur des liquides permet d'avoir une certaine uniformité des cellules sur une lame, contrairement à des prélèvements gynécologiques où l'on travaille sur des étalements. De plus, parmi les différentes cytologies, la cytologie des séreuses fait partie des plus complexes (au niveau de la configuration des cellules et de la diversité des types cellulaires rencontrés). Notre système s'appelle ARCTIC, ce qui signifie Aide à la Recherche en Cytologie par le Tri Informatique Cellulaire. Nous introduisons ici la notion de tri cellulaire qui est intrinsèque à un système de *rescreening*. En effet, la classification (voir paragraphe précédent) permet de déterminer le type d'une cellule et par conséquent de trier les différentes cellules. Cela présente un avantage. Au lieu de simplement déterminer si la cellule est bénigne ou maligne, on détermine en plus son type cellulaire (parmi ceux présentés au paragraphe 1.1.6). Cela permet de disposer d'une information globale sur les cellules présentes sur la lame mais également de connaître, comme pour tout système de *rescreening*, les cellules potentiellement anormales. Le système permet en outre la consultation des caractéristiques attachées à chaque cellule.

ARCTIC est un système permettant donc une forme d'aide au diagnostic. Il vérifie qu'aucune cellule anormale n'est présente sur la lame. Les cellules détectées seront affichées à l'écran suivant leur type cellulaire pour une relecture dans un premier temps par un cytotechnicien (en parallèle du *screening* conventionnel) puis par un pathologiste dans un deuxième temps afin d'établir un diagnostic. Le système aidera donc le pathologiste en vue d'établir un diagnostic en recherchant les éléments rares ou suspects qui auraient pu échapper au *screening* manuel. Il permettra de réduire le nombre de faux-négatifs par la mise en place d'un contrôle qualité (Figure 7).

Le système ARCTIC possède la même architecture que les systèmes de *rescreening* que nous avons présentés. Il est basé sur la segmentation, la caractérisation et la classification des cellules. La segmentation est réalisée par la stratégie de segmentation que nous avons exposée dans la première partie de la thèse. Les cellules sont ensuite caractérisées par les attributs que nous avons présentés dans cette même première partie traitant également de la caractérisation couleur. Les cellules sont donc décrites par des attributs de taille, de forme, de couleur et de texture. La classification se fait par l'architecture MONNA. Les attributs ont été sélectionnés par la méthode *SFFS* (*Sequential Float Forward Selection*) car elle permet d'obtenir le meilleur taux de reconnaissance. Nous n'allons pas décrire à nouveau ces trois étapes séquentielles du système ARCTIC, le lecteur se reportera aux parties concernées. Nous allons cependant décrire les conditions dans lesquelles les images ont été acquises et les conséquences liées à cette phase d'acquisition.

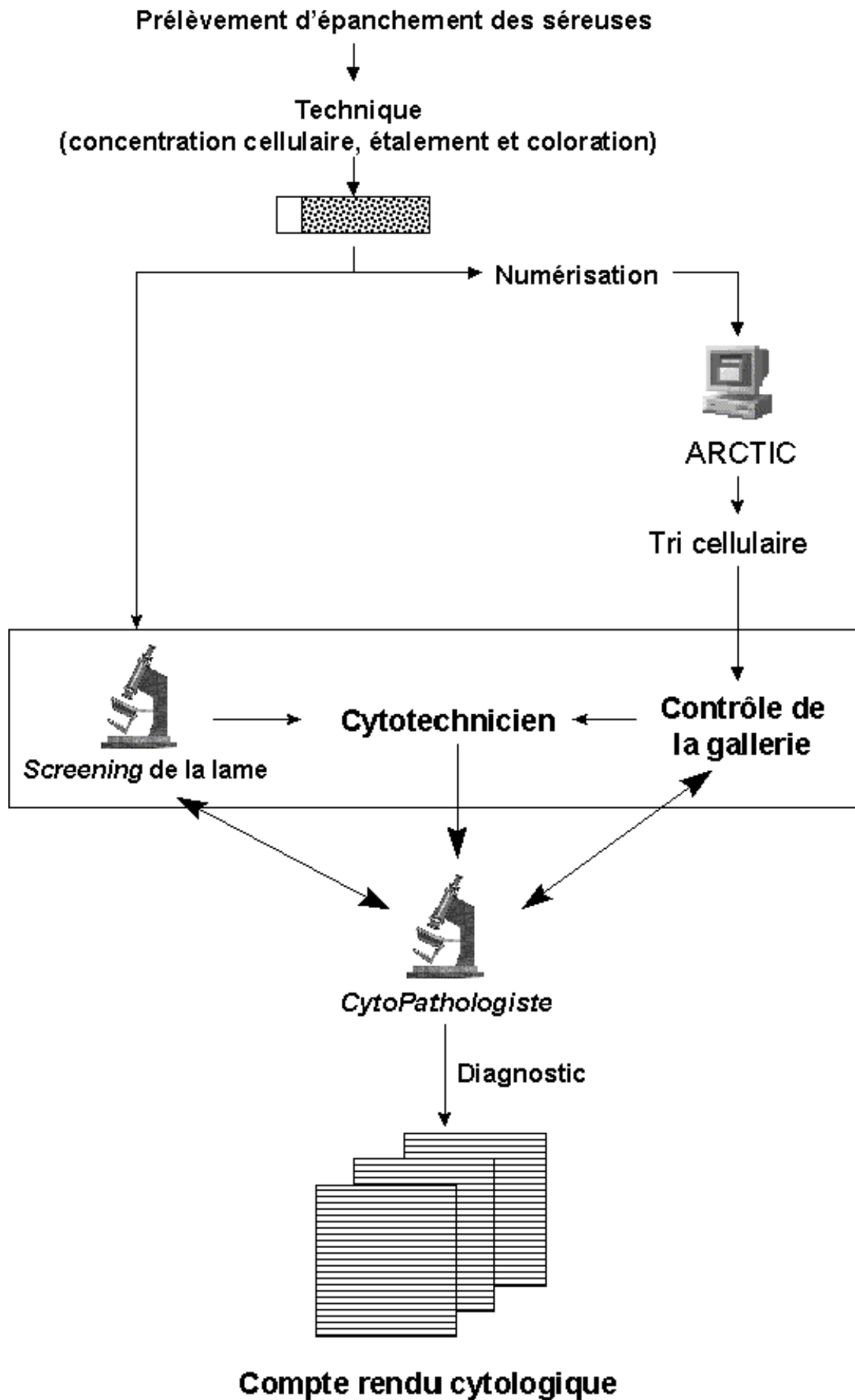


Figure 7. Utilisation d'ARCTIC pour le contrôle de qualité.

2. L'acquisition des images

Pour l'acquisition, nous utilisons un microscope OLYMPUS BX-50 sur lequel est montée une caméra tri-CCD couleur SONY XC-003P reliée à une carte d'acquisition sur un PC (Figure 8).

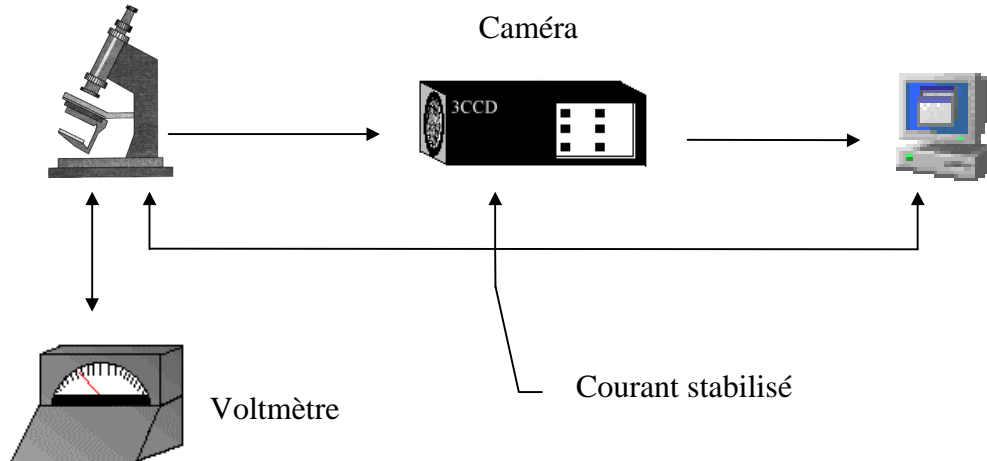


Figure 8. Le système d'acquisition d'images microscopiques.

Pour obtenir une constance dans la tension du courant fourni à la lampe du microscope, nous avons ajouté un voltmètre en parallèle afin de toujours pouvoir contrôler la tension visuellement. Dans notre cas nous gardons cette tension constamment à la valeur 9 Volts (réglée manuellement par une molette du microscope et contrôlée visuellement à l'aide d'un voltmètre), ce qui correspond à une température de couleur de 5500 K. Nous avons choisi ce chiffre en accord avec les spécifications fournies par le constructeur du microscope [Olympus 96] et les spécifications colorimétriques de la CIE.

Afin de minimiser le risque de fluctuations dans l'alimentation en courant de notre système, tous les appareils (microscope, caméra et PC) sont alimentés par un courant stabilisé.

Toutes nos images sont obtenues avec un grossissement et une ouverture de diaphragme constants.

Notre caméra possède un réglage qui permet de la calibrer par rapport à une image blanche. Nous avons fait cet étalonnage par rapport à une lame ne présentant aucun objet (il n'y a rien sur la lame, l'image nous apparaît donc blanche). Cette image a été mémorisée par la caméra pour son calibrage interne. Cela permet d'améliorer la qualité des images car cela réduit les distorsions dues aux capteurs CCD.

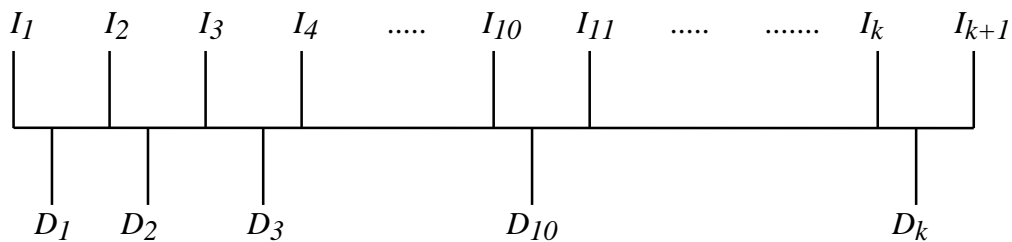
2.1. Résultats expérimentaux sur le système

Afin de bien connaître la réponse de notre système aux différentes perturbations, nous avons procédé à plusieurs tests.

2.1.1. Influence du temps de réponse caméra plus microscope

Pour vérifier la linéarité dans le temps des images obtenues avec notre système, nous avons réalisé la capture d'une même image (représentant un fond blanc uniforme) toutes les cinq minutes pendant deux heures, ceci après avoir allumé simultanément caméra et microscope.

Nous avons ensuite calculé la différence entre deux images successives, cette différence représente ce que nous appellerons « le bruit d'acquisition » d'une image à une autre. Nous avons alors pu quantifier ce bruit grâce aux images des différences et nous avons calculé la moyenne et la variance du bruit d'acquisition (en niveaux de gris) sur chacune des composantes couleurs rouge, vert et bleu.



I_k représente une image au temps $5*k$ minutes et D_k la différence entre l'image I_k et l'image I_{k+1} . Les moyennes et variances sont reportées sur les courbes des figures 9 et 10.

On constate immédiatement que le bruit diminue avec le temps. En fait ceci vient de deux phénomènes :

- La matrice CCD de notre caméra couleur fonctionne à température ambiante et un bruit thermique dû à la chauffe de la caméra s'ajoute à la mesure de l'intensité lumineuse d'un pixel,
- La lampe du microscope chauffe également et l'éclairage de la lampe va croissant au fur et à mesure que le filament chauffe.

On se trouve donc face à un bruit thermique qui est produit par notre système. Ce bruit est dû à la stabilisation en température du système. Au vu des courbes, on s'aperçoit que ce bruit tend à décroître avec le temps. Si l'on modélise l'intensité associée à un pixel dans notre système, on obtient la formule :

$$I_{lue} = I_{exacte} + I_{bruit} + I_{thermique} \quad [\text{Tarel 94,96}]$$

I_{lue} est l'intensité lumineuse que donne le système pour un pixel, I_{exacte} est l'intensité lumineuse qui devrait être donnée par le système si celui-ci était parfait, I_{bruit} est un biais additif constant qui représente le bruit généré par notre système, $I_{thermique}$ est le bruit thermique additif provoqué par l'ensemble caméra et microscope.

Au bout d'environ une heure, $I_{thermique}$ devient négligeable par rapport à I_{bruit} car on observe une stabilité du bruit qui n'est plus que le résultat de l'action de I_{bruit} sur I_{exacte} .

Lorsque l'on fera une acquisition d'image, cela devra se faire au moins une heure après avoir allumé simultanément caméra et microscope, ceci afin de minimiser le bruit thermique. Cela permettra d'éviter d'avoir des erreurs dues à une déviation dans la réponse du système.

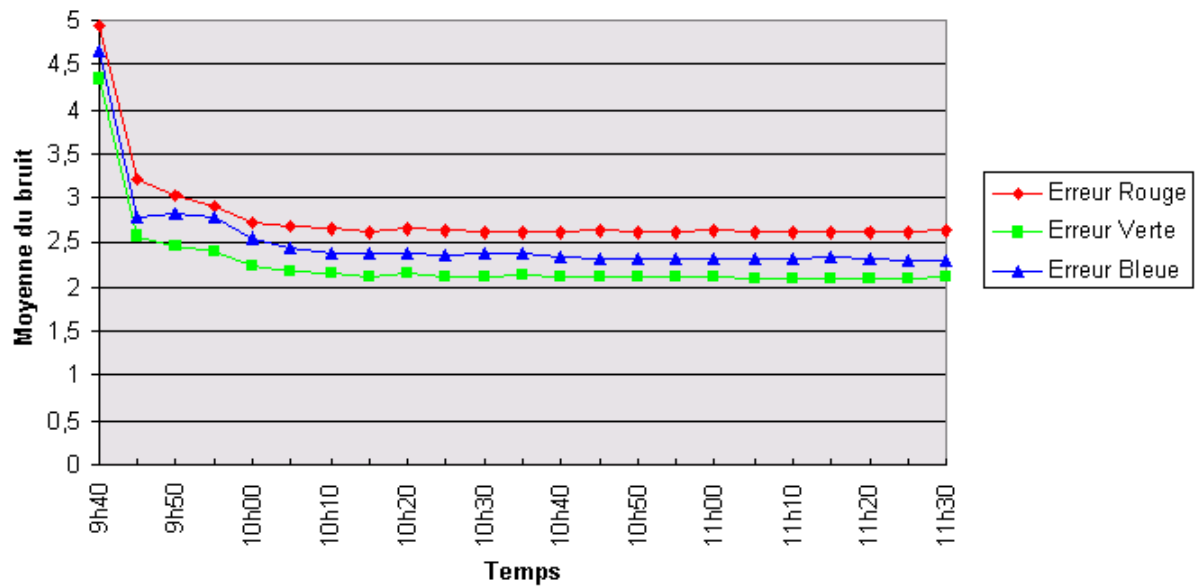


Figure 9. Moyenne du bruit sur les composantes R,V et B en fonction de l'heure d'acquisition.

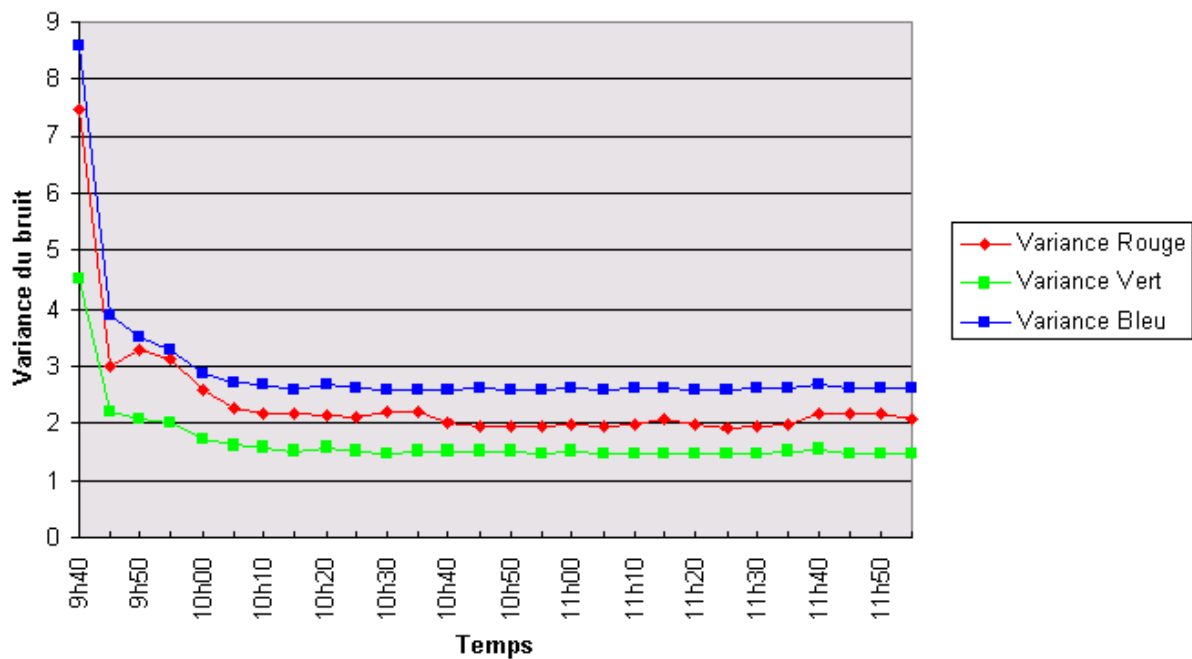


Figure 10. Variance du bruit sur les composantes R,V et B en fonction de l'heure d'acquisition.

2.1.2. Influence de l'environnement

Nous avons procédé à l'acquisition d'images avec plusieurs sources lumineuses parasites : la fenêtre et la lumière artificielle du néon de la pièce. Comme on s'y attendait, chacune de ces sources lumineuses parasites accentue le bruit sur chaque image. En fait cela revient à ajouter à la description précédente du bruit un biais additif $I_{lumière}$. Au vu des

résultats (Figure 11), chaque acquisition doit se faire dans un noir total avec pour seule source lumineuse celle du microscope, ceci afin de minimiser le bruit d'acquisition et de ne pas ajouter d'interférences extérieures.

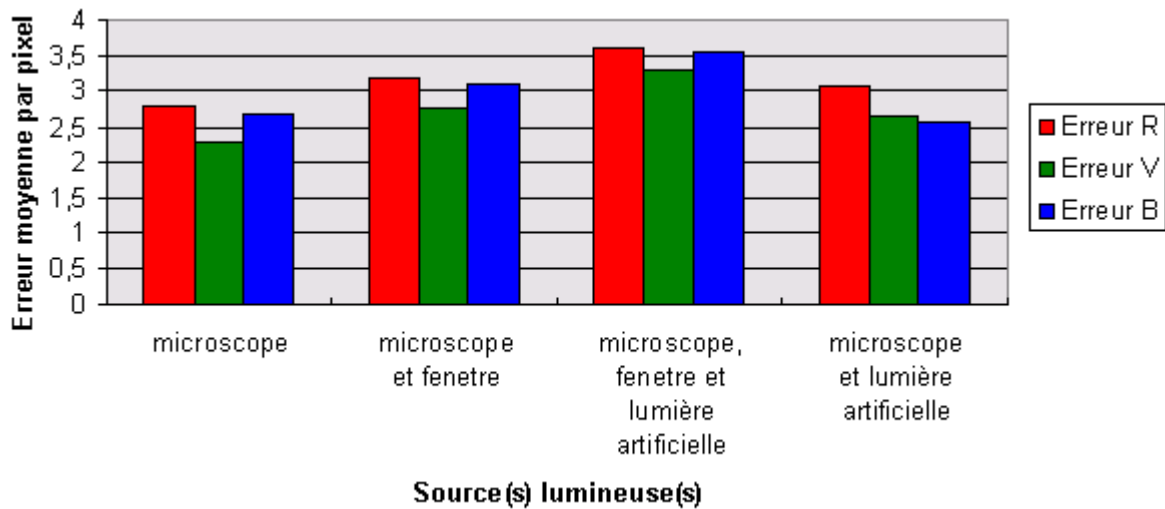


Figure 11. Erreur moyenne sur chaque composante en fonction des sources lumineuses utilisées.

2.2. Détermination du type de bruit du système

2.2.1. Rappels généraux sur le bruit

Les régions formant l'image sont caractérisées par leur intensité moyenne. Les fluctuations autour de cette intensité moyenne peuvent provenir soit du dispositif d'acquisition (caméra, microscope, quantification), soit de la scène elle-même (poussières, rayures...). Elles constituent des perturbations qui affectent la qualité de la segmentation et sont généralement désignées sous le terme de bruit d'image. L'échelle spatiale des fluctuations est relativement faible par rapport aux dimensions des régions. Le bruit est donc plutôt de type haute fréquence. Le bruit d'image est considéré comme étant aléatoire, centré et additif.

La nature plus ou moins impulsionnelle du bruit peut être décrite à l'aide de densités de probabilités [Cocquerez 95]:

$$f(a) = C \cdot e^{-k|a|^\alpha}$$

si $\alpha = \infty$ on a un bruit uniforme,

si $\alpha = 2$ on a un bruit gaussien,

si $\alpha = 1$ on a un bruit exponentiel.

Le caractère impulsionnel du bruit est lié à la probabilité d'avoir des valeurs aberrantes (par exemple d'avoir des intensités s'écartant de plus de trois fois l'écart-type) :

Bruit	Uniforme	Gaussien	Exponentiel
probabilité d'avoir des valeurs aberrantes	0,0	0,003	0,015

2.2.2. Bruit d'acquisition de notre système

Afin de connaître le type de bruit de notre système, nous avons représenté le nombre de pixels bruités en fonction de la valeur en niveaux de gris du bruit. Pour obtenir ces résultats, on procède, comme au paragraphe 2.1, à la différence entre deux images d'un blanc uniforme acquises successivement ; la différence donnant une image caractérisant le bruit d'acquisition. Le bruit d'acquisition a été calculé sur chacune des composantes rouge, vert et bleu de dix images successives.

Les images sur lesquelles nous avons caractérisé le bruit ont bien évidemment été acquises une fois que le système, caméra plus microscope, a atteint sa stabilité thermique (à savoir au bout d'une heure).

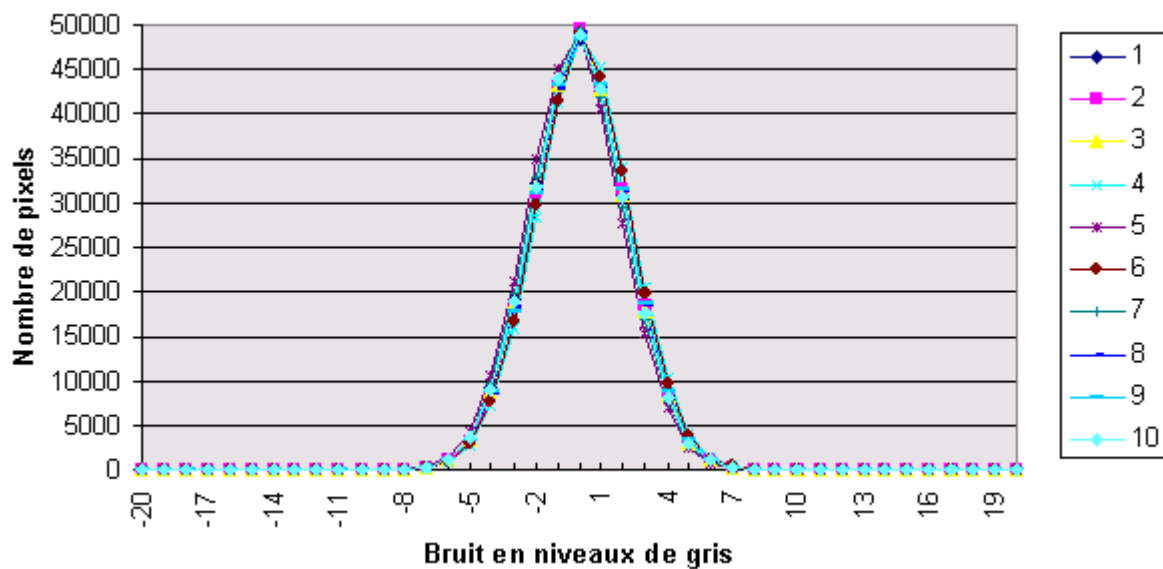


Figure 12. Nombre de pixels en fonction du bruit sur la composante verte.

On remarque que sur chaque composante, le bruit d'acquisition est représenté par une variable aléatoire gaussienne centrée. Le bruit peut varier de 16 niveaux de gris d'un pixel à un autre soit 6,27 % de variation en moyenne. La figure 12 donne le nombre de pixels en fonction du bruit d'acquisition (en niveaux de gris) sur la composante verte. Les courbes obtenues sur les composantes rouge et bleu ont des formes identiques.

Le bruit de notre système est donc conforme au modèle que nous avons évoqué précédemment. Ce qu'il est important de préciser, c'est que ce bruit est toujours de la même forme (car les courbes gaussiennes obtenues à des intervalles de temps différents sont identiques).

2.3. *Élimination du bruit*

Nous avons pu déterminer qu'un certain temps est nécessaire pour obtenir une stabilisation thermique de l'ensemble caméra et microscope. Nous connaissons donc les conditions idéales d'acquisition, cependant les images obtenues sont bruitées, ce bruit étant représenté par une variable aléatoire gaussienne centrée. Nous avons choisi de n'effectuer aucun prétraitement aux images couleur acquises afin de ne pas surcharger l'acquisition qui se

doit d'être rapide. Par contre, lors de la segmentation, nous effectuons des lissages afin d'améliorer et de simplifier les calculs (voir la partie segmentation et caractérisation couleur). Outre le bruit d'acquisition présent sur les images, on peut constater quelques aberrations présentes sur les images. Elles sont dues à la caméra. Afin d'obtenir des images ne présentant pas ce type d'aberrations, nous procédons à une division entre chaque image acquise et une image de fond (une image d'un champ microscopique vide). Les flèches montrent les aberrations présentes sur l'image originale (Figure 13.a). On constate que ces dernières n'apparaissent pas sur l'image corrigée (Figure 13.b). Ceci a plusieurs avantages : les aberrations sont éliminées, le contraste de l'image s'en trouve augmenté et cela permet d'homogénéiser les images entre elles (au niveau de la couleur).

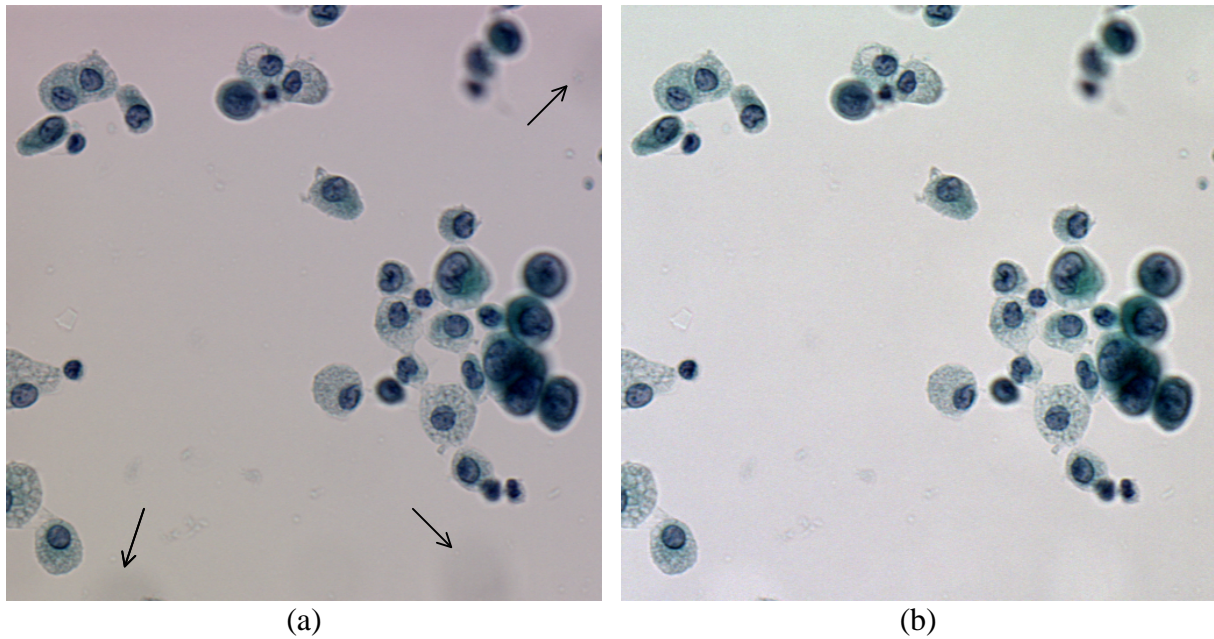


Figure 13. Une image des séreuses avant (a) et après (b) division par une image de fond.

3. Résultats expérimentaux

Nous allons maintenant examiner le fonctionnement du système ARCTIC à travers des résultats expérimentaux. Dans un premier temps, nous avons évalué la segmentation des cellules puis la classification des cellules par l'architecture MONNA.

3.1. Evaluation de la segmentation

L'évaluation d'une segmentation est une tâche difficile. Une inspection visuelle est une méthode simple et facile à implanter, mais étant donné que l'on se base sur le jugement de l'expert, cette méthode est subjective. Une méthode plus objective pour évaluer la précision d'une segmentation consiste à déterminer manuellement les frontières cytoplasmiques et nucléaires de toutes les cellules et à comparer les résultats avec la segmentation en comparant les pixels mal classés. Cette méthode est très fastidieuse et consommatrice de temps. C'est pourquoi, nous avons procédé à une évaluation légèrement différente. Le pathologiste inspecte visuellement les frontières des cellules qui ont été déterminées par la segmentation. Il choisit ensuite les cellules mal segmentées ou celles qui ont été oubliées. Pour celles-ci, il détermine les contours corrects du noyau et/ou du cytoplasme. Cette méthode semi-quantitative est utilisée afin de déterminer le taux de cellules bien segmentées. Toutes les cellules sont examinées par le pathologiste même les amas ou les cellules se chevauchant. Cependant, lorsqu'un expert ne peut pas déterminer de manière fiable les frontières contours de cellules, parce qu'il y a trop de chevauchement entre les cellules, ces dernières sont marquées par le pathologiste comme des cellules non significatives et elles ne sont pas utilisées dans le calcul du taux de segmentation.

Une base de 50 images a été utilisée afin d'évaluer la segmentation. Les cellules segmentées ont été évaluées par trois experts. Une première constatation s'impose : toutes les cellules ont été segmentées et aucune n'a été oubliée. Ceci est très important puisque le système ARCTIC doit effectuer un contrôle qualité. De plus, tous les objets extraits par la stratégie de segmentation sont des cellules (pas de débris ni de globules rouges). Les erreurs de segmentation ont été réparties suivant quatre catégories : un cytoplasme trop grand, un noyau trop grand, un cytoplasme trop petit et un noyau trop petit. Les deux premiers types d'erreurs sont associés à une segmentation majorée et les deux derniers à une segmentation minorée. Pour chaque expert, à l'issue de l'évaluation, le taux de réussite moyen de la segmentation varie entre 94,5% pour les noyaux et 93,0% pour les cytoplastes (Tableau 1).

Numéro de l'expert	Noyaux Corrects	Cytoplastes corrects
1	98,3%	99,0%
2	96,1%	91,5%
3	89,2%	88,7%
Moyenne	94,5%	93,0%

Tableau 1. Taux de segmentation correcte pour les cellules.

Le taux de segmentation correcte des cellules varie de 89,2% à 96,1% pour les noyaux et de 88,7% à 99,0% pour les cytoplastes. Les résultats varient d'un expert à l'autre (environ 10%). Étant donné la difficulté d'évaluation pour des erreurs minimales et pour des cas ambigus, cette différence est considérée comme inévitable. En ce sens, bien qu'une inspection visuelle semi-quantitative permette une mesure plus objective du taux de segmentation

correcte par rapport à une inspection visuelle, l'appréciation du contour des cellules est encore subjective et cela rejoint les limitations du *screening* manuel.

Afin d'apprécier les différences entre les différents types d'erreurs, le taux de cellules mal segmentées est présenté selon le type d'erreur. Ces erreurs sont principalement dues à une segmentation majorée et cela peut être vérifié quel que soit l'expert (Tableau 2).

Numéro de l'expert	Noyaux à segmentation minorée	Noyaux à segmentation majorée	Cytoplasmes à segmentation minorée	Cytoplasmes à segmentation majorée
1	12,5%	87,5%	26,9%	73,1%
2	19,1%	80,9%	37,2%	62,8%
3	27,3%	72,7%	35,1%	64,9%

Tableau 2. Taux d'erreur de la segmentation selon le type d'erreur.

La segmentation majorée des noyaux provient principalement d'amas de cellules dont les noyaux se touchent et qui n'ont pas été séparés. La séparation des noyaux se touchant est basée sur la présence de concavités dans le contour nucléaire des cellules (voir première partie). Or pour certaines configurations où peu de concavités sont présentes, les cellules ne sont pas séparées. 50% des erreurs de segmentation majorée des noyaux sont dues à ce type de configurations. Les autres erreurs proviennent d'une mauvaise mise au point de l'image au niveau des cellules.

La segmentation majorée des cytoplasmes est directement liée à la segmentation majorée des noyaux. Comme les noyaux n'ont pas été séparés, les cytoplasmes ne l'ont pas été non plus. Ce type d'erreur ne pourra être évité que si la segmentation majorée des noyaux a été évitée. La seconde origine de la segmentation majorée des cytoplasmes est liée à la présence de mucus. Ce dernier a la même couleur et texture qu'un cytoplasme et il peut être inclus dans les cytoplasmes segmentés.

La segmentation minorée des cytoplasmes est liée à deux configurations : une non linéarité dans le cytoplasme (un trou ou une vacuole) ou bien plusieurs amas de noyaux se chevauchant sans qu'il y ait de transitions visibles entre les différents cytoplasmes. Dans ce dernier cas, la séparation du cytoplasme profite des positions respectives des noyaux mais ne peut pas tirer parti des transitions entre les cytoplasmes car elles sont peu ou pas apparentes. Les cytoplasmes obtenus ne sont alors pas corrects.

Des exemples d'objets mal segmentés sont donnés par la figure 14.

Les résultats de la segmentation sont donc très satisfaisants. 94,5% des noyaux et 93% des cytoplasmes sont correctement segmentés. Le taux de réussite de la segmentation dépend en grande partie de la population cellulaire présente sur une image. Si beaucoup de cellules isolées sont présentes, la stratégie de segmentation sera très efficace. Mais pour des configurations plus complexes (comme la superposition de noyaux dans les amas), certaines erreurs peuvent apparaître. Les configurations en amas peuvent être extrêmement complexes et leur segmentation complète demande une étude très poussée. L'amélioration de la segmentation ne peut se faire qu'en concevant une méthode spécifique de segmentation des amas. Ceci permettra de corriger la plus grande partie des erreurs. L'autre source d'erreurs provient de la mauvaise mise au point sur une cellule rendant ses contours flous. Les images que nous avons utilisées ont été obtenues par une mise au point globale de toute l'image. Cette

mise au point a été faite manuellement. Une mise au point automatique permettrait d'améliorer la segmentation pour les cellules qui ont été acquises floues [Tomczak 98].

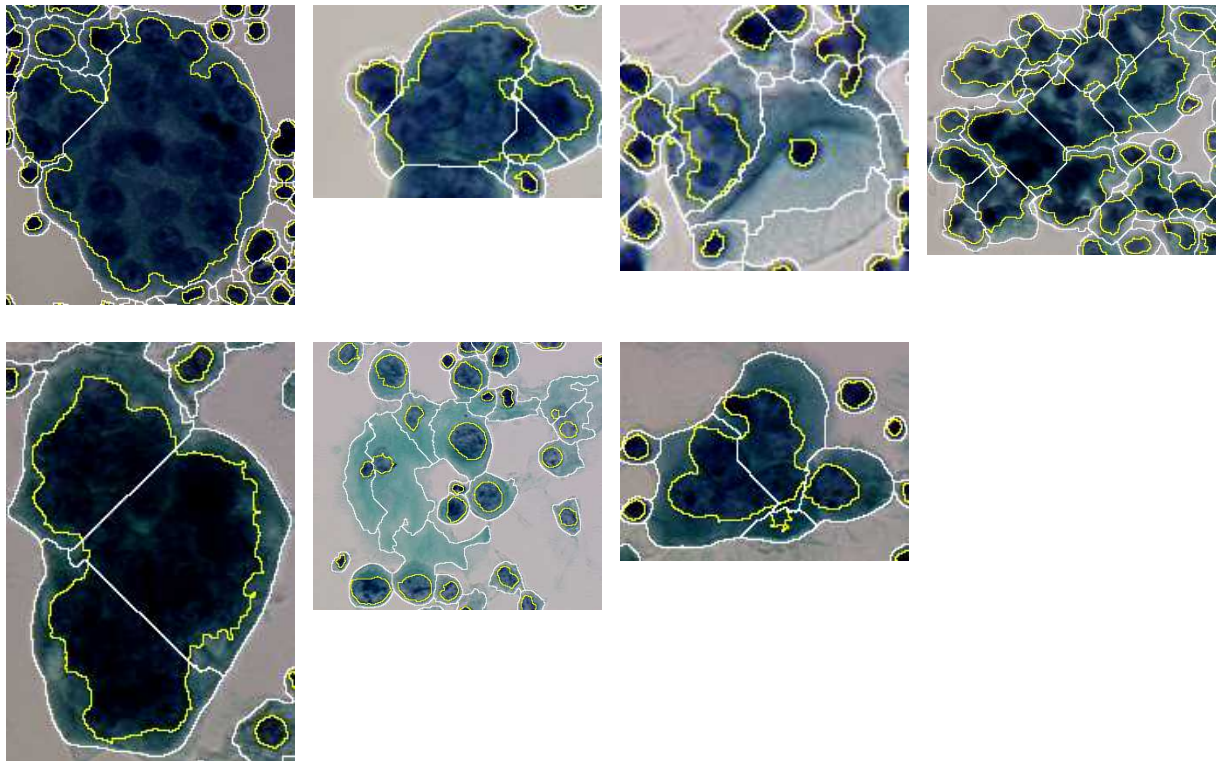


Figure 14. Des erreurs de segmentation.

3.2. Construction d'une base de données

Une fois les cellules segmentées et caractérisées (i.e. leurs attributs ont été calculés), il nous faut constituer une base de données qui va permettre « d'apprendre » au classifieur à reconnaître les cellules. Cette base est construite à partir des cellules segmentées qui proviennent de lames différentes, ceci afin d'avoir une base plus générale.

3.2.1. L'étiquetage par un expert

Tout d'abord, il nous faut définir précisément les classes auxquelles les objets peuvent appartenir. Nous avons répertorié au paragraphe 1.1.6 tous les types cellulaires pouvant être rencontrés. A partir de ces types cellulaires, nous avons déterminé les classes de cellules devant être reconnues par notre système ARCTIC (Tableau 3).

Ces classes peuvent être également regroupées afin d'avoir une vue d'ensemble plus globale de la répartition des cellules, ce qui peut être utile pour un pathologiste.

L'étiquetage des cellules que nous avons mis en place est sensiblement différent des étiquetages habituellement utilisés. Lors d'un étiquetage, l'expert doit décider de l'appartenance d'une cellule à une classe. Or pour certaines cellules, un expert en pathologie ne peut déterminer de façon certaine la classe de celles-ci. L'expert hésite alors entre plusieurs classes. Nous avons pris l'option de ne pas demander à l'expert d'effectuer un choix lorsque ce dernier hésite. L'expert peut donc donner plusieurs classes à une cellule. A l'issue de l'expertise, la base est constituée d'objets qui peuvent avoir une ou plusieurs classes. Les cellules qui n'ont qu'une classe sont considérées comme « sûres » et peuvent être utilisées pour l'apprentissage, par contre les cellules qui ont plusieurs classes sont considérées comme « moins sûres » et ne sont pas utilisées pour l'apprentissage. Cependant, ces cellules peuvent

être utilisées pour tester le classifieur une fois l'apprentissage effectué. Une réponse sera considérée comme exacte si elle est parmi les classes données par l'expert.

Numéro de la classe	Nom de la classe	Regroupement
1	Polynucléaires neutrophiles	Polynucléaires
2	Polynucléaires éosinophiles	
3	Polynucléaires altérés	
4	Lobes de polynucléaire	
5	Lymphocytes non activés	Lymphocytes
6	Lymphocytes activés	
7	Macrophages	Macrophages
8	Sidérophages	Sidérophages
9	Mésothéliales	Mésothéliales
10	Mésothéliales dystrophiques	
11	Mésothéliomes	Cellules anormales
12	Adénocarcinomes	
13	Carcinomes épidermoïdes	
14	Lymphomes	
15	Sarcomes	
16	Mélanomes malins	
17	Mitoses anormales	Mitoses
18	Mitoses normales	

Tableau 3. Les classes de cellules.

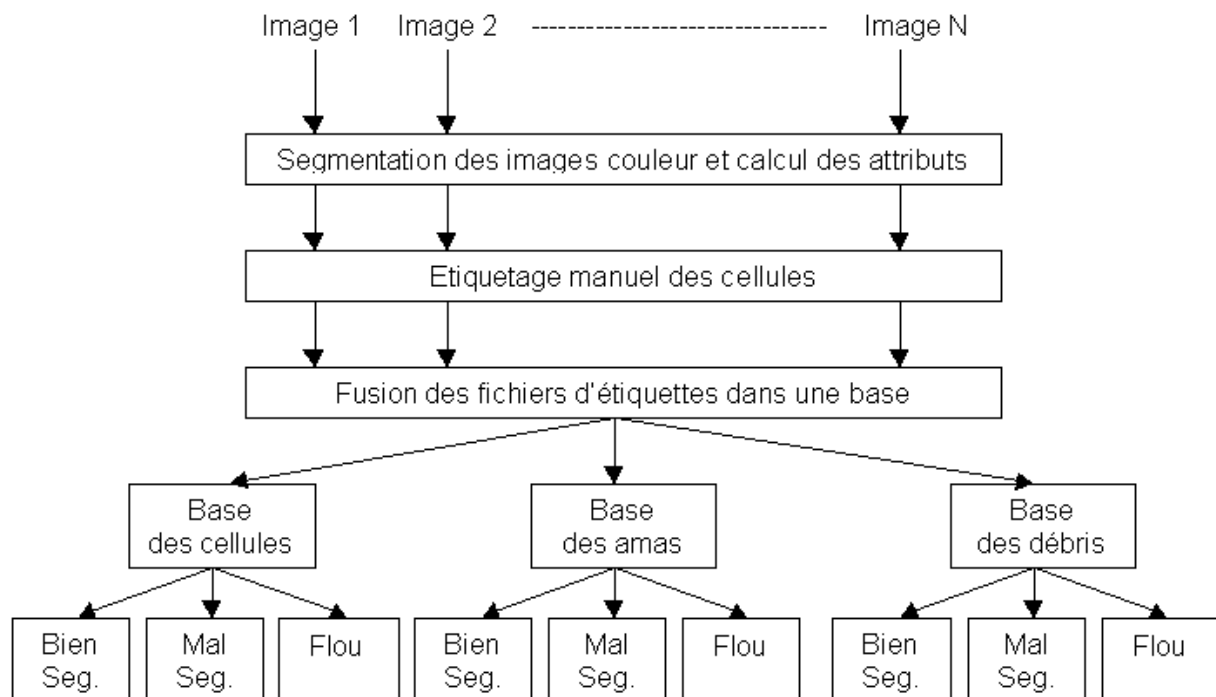


Figure 15. La constitution de bases d'objets par un expert.

De plus, l'expert fournit lors de l'étiquetage des informations sur la cellule ou plutôt sur l'objet qu'il classe. En effet, un objet peut être, à l'issue de la segmentation, soit un amas de cellules, soit une cellule, soit un débris. Ceci doit être pris en compte lors de l'étiquetage. L'expert précise donc à quel type d'objet il a affaire. Mais l'expert précise également si l'objet est flou, mal segmenté ou bien segmenté. Ainsi, nous disposons d'un maximum d'informations sur les cellules, ce qui permet de constituer des bases de données très fiables. L'expert procède à l'étiquetage par une fonctionnalité du logiciel ARCTIC, puis procède à

une vérification systématique de son propre étiquetage en contrôlant une galerie des objets qui ont été classés.

Un expert procède à l'étiquetage de toutes les images de la base d'images (après que celles-ci aient été segmentées et les cellules caractérisées). Un fichier d'étiquetage est généré pour chaque image. Ces fichiers sont fusionnés dans une base qui est divisée en 9 bases différentes (Figure 15). Une première division selon que l'objet soit un amas, une cellule ou un débris et une seconde division de chacune de ces trois précédentes bases selon les indications données sur l'objet (flou, bien ou mal segmenté). Seules les bases qui contiennent des objets bien segmentés peuvent être utilisées. Les autres bases peuvent éventuellement servir de bases de tests pour valider le classifieur utilisé.

3.2.2. Intersection entre plusieurs experts

Les images ayant été étiquetées par plusieurs experts, nous disposons de 9 bases par expert soit 27 bases en tout. Mais seulement 9 d'entre elles (les cellules, amas et débris qui sont bien segmentés) sont utilisées ultérieurement. Afin d'obtenir de comparer les étiquetages des différents experts, nous procédons à une intersection entre les bases des experts. L'intersection permet d'obtenir les objets identiques pour deux experts (classes identiques). L'intersection nous donne également le taux d'objets identiques entre les experts. Nous avons examiné tout d'abord le taux d'objets communs qui ont été classés comme amas, cellule isolée et débris par deux experts. Ces taux sont supérieurs à 90% (Tableau 4), on peut donc dire que les experts sont en accord à ce niveau. Nous avons procédé à l'intersection des bases de cellules bien segmentées. Le taux de cellules à classes identiques entre deux experts est relativement peu élevé (entre 54% et 65%). Les experts ont donc des difficultés à s'accorder entre eux. Ceci est considéré comme normal pour deux raisons. Premièrement, l'évaluation des cellules des cellules et le nombre de cellules à évaluer dans la base est élevé (environ 5000). Encore une fois, nous rejoignons ici les limitations du *screening* conventionnel. Cela se vérifie également pour les bases des amas bien segmentés et des débris bien segmentés. Deuxièmement, les experts ont des difficultés à exprimer leur connaissance. Elle peut être vue comme une connaissance « compilée », c'est-à-dire que les experts la possèdent mais ils ont du mal à l'exprimer. Ceci se vérifie d'autant plus qu'il ne sont pas habitués à déterminer précisément quel est la classe d'une cellule lors du screening (l'évaluation de la lame est plus globale). Si bien que la mise au point de la base les amène à prendre des décisions difficiles, ce qui explique les différences entre les experts.

Amas, Cellule ou Débris	Taux
$1 \cap 2$	92,85%
$1 \cap 3$	96,09%
$2 \cap 3$	93,93%

Cellules B.S	Taux	Amas B.S	Taux	Débris B.S	Taux
$1 \cap 2$	65,03%	$1 \cap 2$	82,73%	$1 \cap 2$	66,02%
$1 \cap 3$	62,08%	$1 \cap 3$	84,33%	$1 \cap 3$	52,47%
$2 \cap 3$	54,63%	$2 \cap 3$	53,01%	$2 \cap 3$	87,13%

Tableau 4. Les taux de concordance entre les experts suivant la base utilisée.

Finalement, nous disposons de trois bases respectivement pour les amas, les cellules et les débris qui sont bien segmentés.

3.3. Classification des objets

Pour classer les objets d'une image, il nous faut élaborer une stratégie de classification des objets de l'image. Tout d'abord, nous avons construit un arbre de décision (par C4.5 [Quinlan 93]) qui détermine si la cellule est un amas, une cellule ou un débris. Le taux d'erreur de cet arbre est de 1,3%. Les erreurs proviennent de petites cellules qui ont été classées comme débris et de très grandes cellules qui ont été classées comme amas. Ces erreurs ne sont pas très importantes du point de vue cytologique car les petites cellules ne sont pas importantes et les grandes cellules sont généralement suspectes (l'expert devra revoir les amas obtenus car cela est généralement suspect).

Les cellules isolées sont ensuite classées en utilisant l'architecture de réseaux de neurones MONNA avec une sélection d'attributs par la méthode *SFFS*. Dans la deuxième partie de la thèse, nous avons utilisé la base « serous » qui est la base des cellules bien segmentées obtenue par l'expert numéro 1 (appelons-la *B1*). Les tests de cette deuxième partie qui ont été effectués ne portant que sur la base *B1*, les résultats ne sont pas comparables avec ceux que nous allons exposer ici car la base de test est différente. Pour notre classification finale, nous avons utilisé l'union des bases (des cellules bien segmentées) entre les experts 1 et 2 (car la base de l'expert 3 ne contient pas assez de cellules), nous l'appellerons *B2*.

Nous avons effectué l'apprentissage de notre architecture en utilisant la base *B2*, nous rappelons ici que cette base est divisée en trois bases : une base de test, une base d'apprentissage et une base de validation. Le taux de reconnaissance global de l'architecture MONNA en utilisant *B2* pour l'apprentissage avec sélection de caractéristiques par la méthode *SFFS* est de 84,55% pour les cellules à une classe.

Nous n'avons fait figurer dans la matrice de confusion de l'architecture MONNA (avec la base *B2*) que les classes pour lesquelles des cellules sont présentes dans la base d'apprentissage. En effet certaines classes ne sont pas représentées dans la base car ce sont des cellules rares. Le taux de reconnaissance du classifieur MONNA sur les cellules qui ont plusieurs classes possibles est de 78,45%, ce résultat est moins élevé que pour les cellules à une classe, mais il est supérieur aux taux de concordance entre experts. Si l'on s'attarde sur la matrice de confusion, on remarque que les principales erreurs se situent au niveau de :

- la distinction entre les différents types de polynucléaires (classes 1,2,3,4),
- la distinction entre les lymphocytes non activés (n°5), les lymphocytes activés (n°6) et les macrophages (n°7),
- la distinction entre les macrophages (n°7) et les mésothéliales (n°9).

Ces problèmes sont également rencontrés par les experts lors de l'évaluation des cellules. La classe macrophage est très hétérogène et sa reconnaissance est difficile. Un travail d'extension de la base reste à effectuer. Ceci afin d'équilibrer et d'étoffer le nombre de cellules de chaque classe. En effet certaines classes sont peu représentées et donc peu apprises par l'architecture MONNA, ce qui ne permet pas de donner des résultats assez significatifs pour ces classes (les mitoses par exemple).

Classe souhaitée

	1	2	3	4	5	6	7	9	10	11	12	17
1	54	4	6	2	3	1	2	0	0	0	0	0
2	1	52	0	0	1	0	2	0	0	0	0	0
3	17	0	20	1	3	0	3	0	0	0	0	0
4	4	1	1	47	0	0	0	0	0	0	0	0
5	3	1	0	0	135	10	2	0	0	0	0	0
6	4	0	0	0	9	55	4	0	0	0	0	0
7	0	0	0	1	2	4	111	9	0	0	0	0
9	0	0	0	0	0	0	6	109	0	0	5	0
10	0	0	0	0	0	0	0	3	3	0	2	0
11	0	0	0	0	0	0	1	0	0	4	1	0
12	0	0	0	0	0	0	0	1	0	0	79	0
17	0	0	0	0	0	0	0	1	0	0	2	4

Classe obtenue

Tableau 5. La matrice de confusion.

Classes	Polynucléaires classes 1,2,3,4	Lymphocytes classes 5,6	Macrophages ou Mésothéliales classes 7,9	Anormales Classes 10 à 17
Polynucléaires	210	8	7	0
Lymphocytes	8	209	6	0
Macrophages ou Mésothéliales	1	6	235	5
Anormales	0	0	6	95

Tableau 6. La matrice de confusion du Tableau 5 après regroupement

Si l'on se place dans l'optique de notre système, à savoir la détection de cellules anormales ou suspectes, notre système est très satisfaisant (Tableau 6) car 94,5% des cellules anormales et 99% des cellules normales sont reconnues, ce qui est largement supérieur au taux de réussite d'un expert. Etant donné que notre système opère en *post-screening*, on peut penser, au vu des résultats obtenus, qu'une cellule anormale omise par le cytotechnicien sera détectée par notre système.

Les taux de classification de l'architecture MONNA sur les cellules floues est de 35,1% pour les cellules à une classe et de 46,3% pour les cellules à plusieurs classes. De même, pour les cellules mal segmentées, le taux de classification est de 41,1% pour les cellules à une classe et de 43,2% pour les cellules à plusieurs classes. Etant donné que la majorité des erreurs de segmentation proviennent d'une mise au point imprécise des cellules, le taux de classification ne peut être amélioré qu'avec une mise au point automatique.

4. Fonctionnalités du système ARCTIC

Le système ARCTIC est un logiciel réalisé en Visual C++[®] sous Windows NT. Les opérations de traitement d'images sont indépendantes de la gestion du fenêtrage et peuvent être incorporées à tout autre logiciel. Nous utilisons la librairie de traitement d'images PANDORE [Clouard 97] à laquelle nous avons ajouté la gestion des images couleur. Cette librairie permet de nombreux traitements sur plusieurs types de données : des images 2D en niveaux de gris et en couleur, des images 3D, et également des graphes [Angot 99]. Le logiciel ARCTIC permet de charger des images en couleur et en niveaux de gris. Les opérations de consultation du contenu de l'image telles que l'histogramme, le profil, le zoom peuvent être utilisées (Figure 16).

Il permet de réaliser la stratégie de segmentation ainsi que l'évaluation des résultats grâce à un outil permettant de dessiner le contour des cellules (Figure 17). L'expert peut sélectionner une cellule et procéder à la correction de la segmentation du noyau et/ou du cytoplasme.

Une fois les images segmentées, le calcul des attributs des cellules peut être réalisé et ces attributs peuvent être consultés pour une image. L'expert peut consulter les valeurs des attributs pour une cellule ainsi que les valeurs d'un attribut pour toutes les cellules (Figure 18.a). Le fichier contenant les caractéristiques peut être sauvé dans un format texte permettant une exportation vers un tableur (Figure 18.b).

L'étiquetage des cellules présentes sur une image est réalisé grâce à une boîte de dialogue. Les cellules sont présentées à l'expert (Figure 19.a) qui précise le type cellulaire, le type d'objet (cellule isolée, amas ou débris) ainsi que des indications relatives à la cellule (si elle est floue ou mal segmentée). On notera que l'étiquetage est différent pour chacun des experts car ils doivent s'identifier avant d'étiqueter une image (Figure 19.b).

Lorsqu'une base a été réalisée, c'est-à-dire que les cellules ont été étiquetées, l'expert peut visualiser cette base. Le logiciel ARCTIC lui présente une galerie d'images selon les types cellulaires (Figure 20.b), chacune de ces images correspondant à une cellule de la base. L'expert peut alors réviser la base et modifier le type cellulaire des cellules (Figure 20.a). Il peut également visualiser les répartitions des cellules selon les types cellulaires (Figure 20.c). Cette visualisation d'une galerie d'images est également utilisée pour consulter le résultat de la classification des cellules d'une image ne faisant pas partie de la base, ce qui est le but principal de notre logiciel, c'est-à-dire visualiser des cellules anormales qui ont été repérées par le système.

Le logiciel intègre également des fonctionnalités nécessaires à l'acquisition d'images provenant d'un microscope : le pilotage d'une platine motorisée (MC2000 de chez Märzhäuser [Huet 98]) et d'une carte d'acquisition (MuTech MV1300).

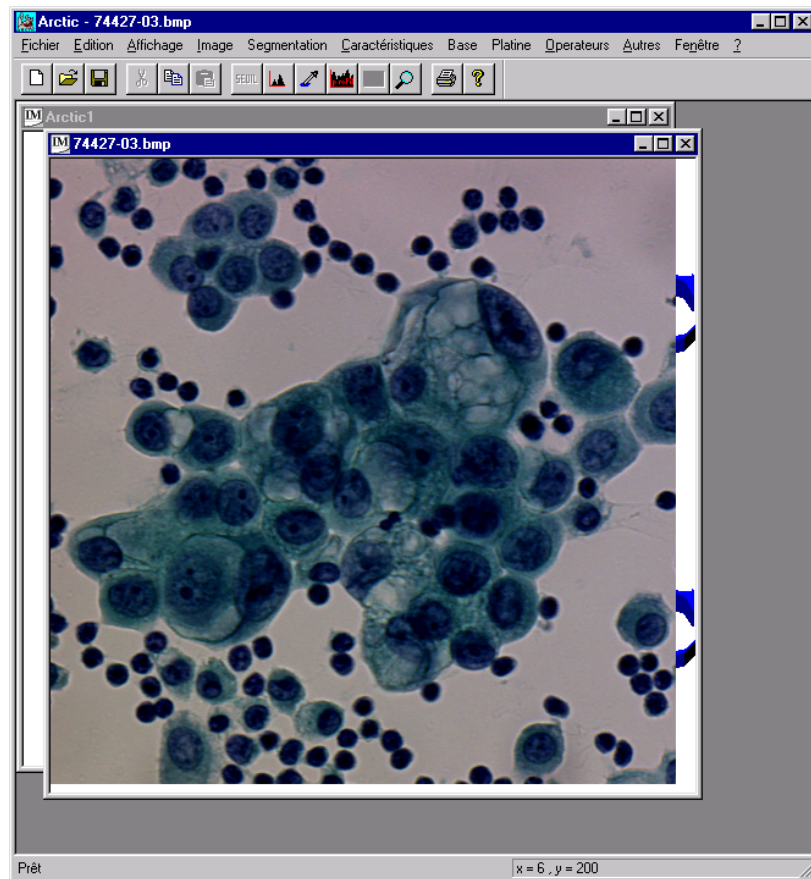
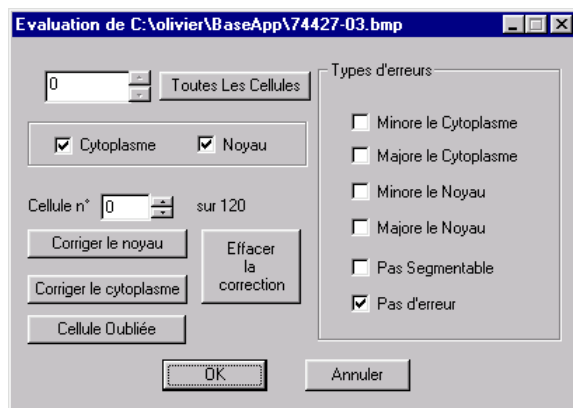


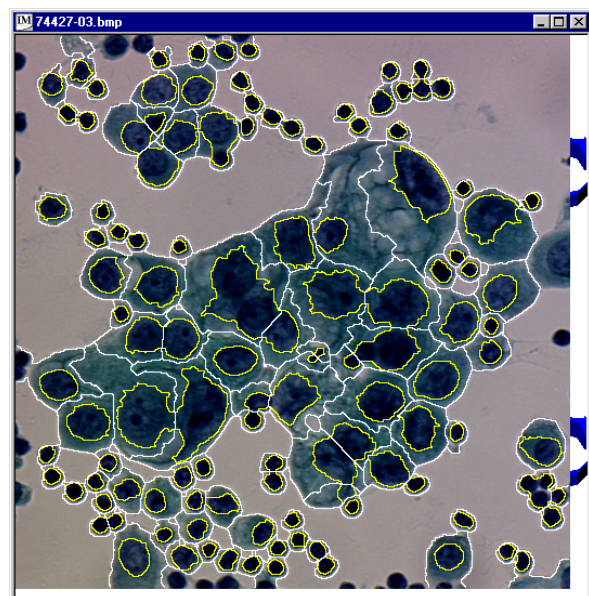
Figure 16. Le logiciel ARCTIC.



(a)

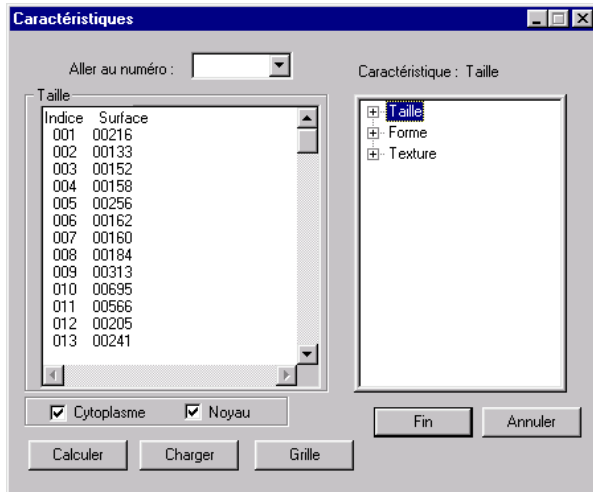
(a) l'évaluation de la segmentation,

(b) les contours obtenus pour les cellules sur une image.

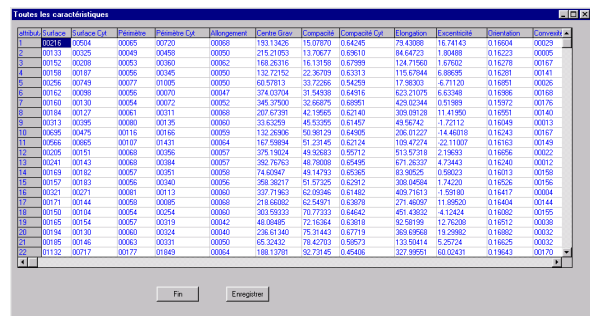


(b)

Figure 17. Le processus d'évaluation de la segmentation d'une image.

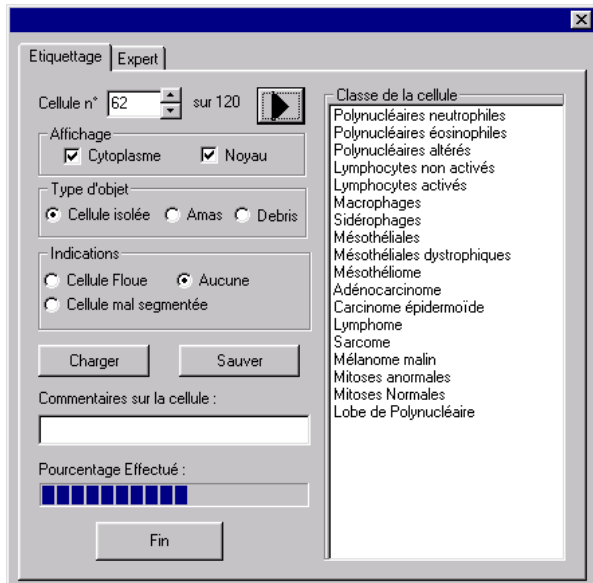


(a) la consultation des attributs des cellules.

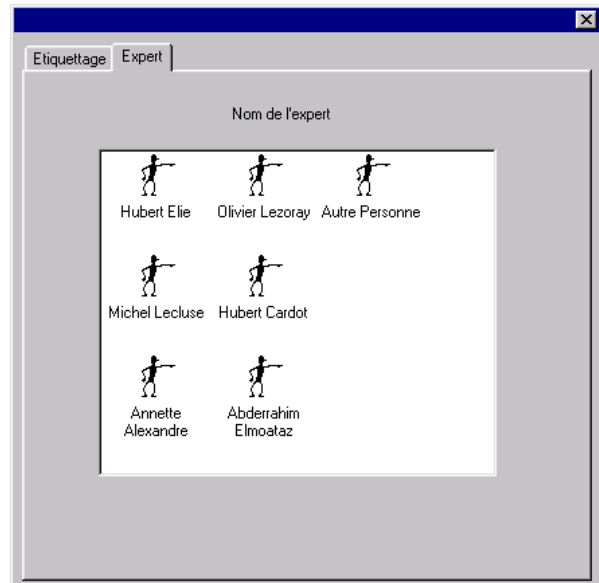


(b) la sauvegarde des attributs pour exportation vers un tableur.

Figure 18. La consultation des caractéristiques des cellules d'une image.



(a) l'étiquetage des cellules.



(b) la spécification de l'expert.

Figure 19. Les boîtes de dialogue d'étiquetage d'une image de cellules.

(a)

(b)

(c)

Classe	Nombre de cellules	Pourcentage
Polynucléaires neutrophiles	1	0.8%
Lymphocytes non activés	57	47.5%
Lymphocytes activés	8	6.7%
Macrophages	11	9.2%
Mésothéliales	10	8.3%
Mésothéliales dystrophiques	1	0.8%
Adénocarcinome	23	19.2%
Lobe de Polynucléaire	2	1.7%
Mésothéliales dystrophiques ou Adénocarcinome	1	0.8%
Amas de Lymphocytes non activés	1	0.8%
Amas de Adénocarcinome	1	0.8%
Amas de Lymphocytes non activés et Mésothéliales	1	0.8%
Amas de Lymphocytes non activés et Adénocarcinome	1	0.8%
Débris	2	1.7%
Total	120	100.0%

- (a) la consultation de la base,
 (b) les imagettes des cellules,
 (c) les proportions des cellules.

Figure 20. La consultation d'une base de cellules.

5. Conclusion

Dans cette troisième partie de la thèse, nous nous sommes intéressés au problème du tri cellulaire en vue d'une aide au *screening*. L'objectif est de développer un système permettant d'extraire les cellules d'une image couleur de la cytologie des séreuses puis de les caractériser et enfin de les classer dans un des types cellulaires possibles. Nous avons utilisé les méthodes et algorithmes qui ont été développés dans les deux premières parties de la thèse. Leur utilisation s'est révélée efficace puisque d'une part les cellules sont globalement bien segmentées : 93% des noyaux et 94,5% des cytoplasmes. D'autre part, la caractérisation couleur des cellules et l'architecture MONNA permettent une bonne reconnaissance des cellules : 94,5% des cellules anormales et 99% des cellules normales sont reconnues, ce qui est largement supérieur au taux de réussite d'un expert. L'élaboration de notre système est donc réussie, cependant certains points restent à approfondir. Une mise au point automatique permettrait d'améliorer de façon significative la segmentation ainsi que la reconnaissance des cellules. De plus, la base doit être complétée par davantage de cellules pour améliorer l'apprentissage et équilibrer les proportions des différentes classes de cellules étant données les grandes variations d'expertise entre les pathologistes dues aux classes difficiles à reconnaître qui sont peu représentées.

6. Références

- [Bartels 87] Bartels P.H., Graham A., Kuhn W., Wied G.L., « Knowledge engineering in quantitative histopathology », *Applied Optics*, Vol. 26, pp. 3330, 1987.
- [Beale 1860] Beale L.S., « Examination of sputum from a case of cancer of the pharynx and adjacent parts », *Arch Med*, Vol. 22, pp. 44, 1860.
- [Boon 93] Boon M.E., Kok L.P., « Neural network processing can provide means to catch errors that slip through human screening of pap smears », *Diagnostic Cytopathology*, Vol. 9, n° 4, 1993.
- [Clouard 97] Clouard R., Elmoataz A., Angot F., Lezoray O., « *Pandore : une bibliothèque et un environnement de programmation d'opérateurs de traitement d'images* », Rapport interne du GREYC, Caen, France, Mars 1997.
- [Cocquerez 95] Cocquerez J-P., Philipp S. « *Analyse d'images : filtrage et segmentation* », éditions Masson, 1995.
- [Derder 95] Derder A., « *Système de support et de supervision des activités en cytologie* », Thèse de l'Institut National de Polytechnique de Grenoble, Thèse IMAG, 1995.
- [Donné 1845] Donné A., « *Cours de microscopie élémentaire des études médicales d'anatomie microscopique et physiologie des fluides* », J.B Baillière, Paris, 1845.
- [Dufour 04] Dufour H.M., « Meningité sarcomateuse diffuse avec envahissement de la moelle et des racines : cytologie positive et spéciale du liquide céphalo-raphidien », *Rev. Neurol. Soc. Neurol.*, Vol. 12, pp. 204, 1904.
- [Dytch 87] Dytch H.E., Bibbo M., Bartels P.H., Puls J.H., Wied G.L., « A PC-based system for the objective analysis of histologic specimens through quantitative contextual karyometry », *Applied Optics*, Vol. 26, n° 16, pp. 3270, 1987.
- [ENSTAC 87] « *Cours de cyto-pathologie* », Ecole nationale supérieure de techniques anatomo et cyto-pathologiques, Université de Paris VII, 1987.
- [Erhardt 80] Erhardt R., Reinhart E.R., Shipf W., Bloss W.H., « FAZYTAN : a system for fast automated cell segmentation, cell image analysis and feature extraction, based on T-V image pickup and parallel processing », *Analytical and Quantitative Cytology and Histology*, Vol. 2, pp. 25-40, 1980.
- [Hall 87] Hall T.L., Keyani D.H.L., Rosenthal D.L., « Microcomputer-based image processing workstations for cytology », *Applied Optics*, Vol. 26, n° 16, pp. 3266, 1987.
- [Huet 98] Huet J., « *Acquisition automatique d'images sur un microscope à platine multilame par référence à une lame numérisée* », Rapport de Stage de DESS Applications Documentaires et Images, Université de Caen, 1998.

[Kasdan 87] Kasdan H.L., « High performance pathology workstation using an automated multispectral microscope », *Applied Optics*, Vol. 26, n° 16, pp. 3295, 1987.

[Knesel 96] Knesel E.A., « Roche Image Analysis Systems, Inc. » *Acta Cytologica*, Vol. 40, n° 1, pp. 60-66, 1996.

[Koss 89] Koss L.G., « The Papanicolaou test for cervical cancer detection, a triumph and a tragedy », *J. Am. Med. Assoc.*, Vol. 261, pp.737-743, 1989.

[Koss 93] Koss L.G., « Cervical Pap smear », *Cancer*, 1993.

[Koss 94] Koss L.G., « Evaluation of PAPNET cytologic screening system for quality control of cervical smears », *American Journal of clinical pathology*, Vol. 101, n° 2, 1994.

[Lebert 1845] Lebert H., « *Physiologie et pathologie ou recherches cliniques, expérimentales et microscopiques* », J.B Baillière, Paris, 1845.

[Mango 93] Mango L.J., « Aide au dépistage du cancer du col par ordinateur à réseaux neuraux », *Cancer Letters*,1993.

[Marissi 09] Marissi G., « Ueber die diagnose des magencarcinomas auf Grund der Cytologie des Spuelwasser », *Arch Verdauungskrankh*, Vol. 15, pp. 251, 1909.

[Menard 93] Menard M., « *Développement d'algorithmes et d'un système de traitement associé, pour l'aide au diagnostic d'images de cellules épithéliales en vue d'examens en routine clinique* », Thèse de l'Université de Poitiers, 1993.

[Mueller 1838] Mueller J., « Ueber den feineren Bau und die Formen der krankhaften geshwuelste », *G Reimer*, Berlin 1838.

[Nordin 89] Nordin B., « *The development of an automated prescreener for the early detection of cervical cancer: Algorithms and Implementation* », Ph.D. Thesis, Upsalla University, Sweden,1989.

[Olympus 96] Manuel Olympus, « *How to Improve Photography Through The Microscope* », Japon, 1996.

[Papanicolaou 28] Papanicolaou G.N., « New cancer diagnosis », *Proc. third race betterment conference*, pp. 528, 1928.

[Papanicolaou 47] Papanicolaou G.N., Traut H.F., « Diagnosis of Uterine Cancer by the Vaginal Smear », *Commonwealth Fund*, NewYork, 1947.

[Papanicolaou 54] Papanicolaou G.N., « *Atlas of Exfoliative Cytology* », Harvard University Press, Cambridge, Massachusetts,1954.

[Patten 96] Patten S.F., Lee J.S.J., Nelson A.C., « Neopath AutoPap 300 Automatic Pap Screener System », *Acta Cytologica*, Vol. 40, n°1, pp. 46-52, 1996.

- [Ploem 87] Ploem J.S., Tanke H.J., « Introduction to fluorescence microscopy », *Roy. Micr. Soc.*, Microscopy handbooks, Oxford, University Press, 1987.
- [Pouchet 1847] Pouchet F.A., « *Théorie positive de l'ovulation spontanée et de la fécondation des mamifères et de l'espèce humaine* », J.B Baillière, Paris, 1847.
- [Prewitt 66] Prewitt J.M.S., Mendelsohn M.L., « The analysis of cell images », *Ann. N.Y. Acad. Sc.*, Vol. 128, pp. 2035, 1966.
- [Quinlan 93] Quinlan J.R., « Programs for machine learning », Morgann Kauffman, San Mateo, CA, 1993.
- [Revision 67] « *Revision of laboratory requisitions for the Medicare, Medicaid and clinical laboratories improvement act of 1967 (CLIA'67) programs* », Federal Register, Vol. 5, pp. 9600, 1967.
- [Rosenthal 96] Rosenthal D.L., Acosta D., Peters R.K., « Computer assisted rescreening of clinically important false negative cervical smears using the papnet testing system », *Acta Cytologica*, Vol. 40, n°1, pp.120-126, 1996.
- [Sanders 1864] Sanders W.R., « Cancer of the bladder : fragments forming urethral plugs discharged in the urine », *Med. J.*, Edinburgh, Vol. 1, pp. 273, 1864.
- [Sherman 94] Sherman M.E., « Papnet analysis of reported negative smears preceding the diagnosis of a high grad squamous intraepithelial lesion or carcinoma », *Modern Pathology*, Vol. 7, n° 5, pp. 578, 1994.
- [Tanaka77] Tanaka N., Ideka H., Ueno T., Watanabe S., Imasato Y., Kashida R., « Fundamental study of cyto-screening for uterine cancer : new system for automated apparatus (CYBEST) utilizing pattern recognition method », *Acta Cytologica*, Vol. 21, pp. 85, 1977.
- [Tanaka 87] Tanaka N., Ideka H., Ueno T., Mukawa A., Watanabe S., Okamoto K., Hosoi S., Tsunekawa S., « Automated cytologic screening system (CYBEST model 4) an integrated image cytometry system », *Applied Optics*, Vol. 26, n° 16, pp. 3301-3307, 1987.
- [Takahashi 83] Takahashi M., « *Atlas en couleur de cytologie du cancer* », deuxième édition, VIGOT, 1983.
- [Tarel 96] TAREL Jean-Philippe, « Une méthode de calibration radiométrique de caméra à focale variable », *Actes du 10^{ème} congrès AFCET, Reconnaissance des Formes et Intelligence Artificielle*, Rennes, Janvier 1996.
- [Tarel 94] TAREL Jean-Philippe, GAGALOWICZ André, « Calibration de caméra à base d'ellipses », *Actes du 9^{ème} congrès AFCET, Reconnaissance des Formes et Intelligence Artificielle*, Paris, Janvier 1996.
- [Tolles 55] Tolles W.E., « The cytoanalyser : an example of physics in medical research », *Trans. N.Y Acad. Sc.*, Vol. 17, pp. 250, 1955.

[Tomczak 98] Tomczak R., Bonton P., « Etude d'un algorithme de mise au point de l'image : application à l'observation d'échantillons sous microscope », *Actes de RFIA '98*, 1998.

[Tucker 82] Tucker J.M., Shippey G.A., Farrow S., Bayley R., Stark M., « CERVIFIP : an automated cytologic prescreening system using a linear diode array », *Analytical Quantitative Cytology*, Vol. 4, pp. 160, 1982.

[Wied 68] Wied G.L., Bartels P.H., Bart G.F., Oldfield D.G., « Taxinomic intra-cellular Analytic System (TICAS) for cell identification », *ActaCytologica*, Vol. 12, pp. 180-204, 1968.

[Zanisher 82] Zanisher D.J., Oud P.S., Raaijmakers M.C.T., Vooijs G.P., Van de Walle R.T., « the BioPEPR system for cervical cytology automation », *Analytical Quantitative cytology*, Vol. 4, pp. 164, 1982.

Conclusion et perspectives

Conclusion et perspectives

Dans ce travail, nous avons étudié les différentes étapes utiles à l'élaboration d'un système de vision par ordinateur : la segmentation d'images, la caractérisation d'objets et la classification de données. Dans les deux premières parties de la thèse, nous avons présenté des méthodes générales qui peuvent être adaptées à différents problèmes de vision par ordinateur. Nous avons développé des techniques originales permettant de réaliser chacune de ces étapes. Premièrement, la couleur est une information importante dans les images et nous avons cherché à exploiter les possibilités liées à l'information couleur, ceci tant au niveau de la représentation de la couleur qu'au niveau de son intégration dans des techniques de traitement d'images. Deuxièmement, la classification est une étape importante dans un système de vision. Nous avons conçu à cet égard une architecture de réseaux de neurones permettant d'améliorer la classification des données. Troisièmement, nous avons élaboré un système d'aide au *screening* par le tri informatique cellulaire.

La première partie présente la segmentation et la caractérisation d'images couleur. Nous avons présenté une méthode de segmentation d'images couleur de microscopie par une ligne de partage des eaux couleur utilisant des informations locales et globales. Afin de pouvoir adapter les points critiques de la méthode, nous avons défini une stratégie d'optimisation qui permet de définir la façon de les aborder et de les optimiser. Nous nous sommes plus particulièrement intéressés au choix de l'espace couleur et à l'obtention des marqueurs. Nous avons vu que le choix de l'espace couleur est un point important. Nous avons défini une stratégie de choix de l'espace couleur basée sur une mesure du contraste couleur. Ceci permet d'obtenir l'espace couleur le mieux approprié pour notre méthode de segmentation couleur en fonction d'informations a priori. Ces dernières sont également utilisées pour l'extraction des marqueurs. Nous avons montré, à travers deux applications, que notre méthode permet d'extraire des objets pour des images où la quantité d'informations a priori est élevée. Cette méthode est donc bien adaptée à la segmentation d'images de microscopie. Nous avons pu mettre en place une stratégie complète de segmentation d'images de la cytologie des séreuses. Nous avons défini également des critères quantitatifs permettant de caractériser les objets extraits lors d'une segmentation. Nous avons plus particulièrement approfondi les attributs permettant de caractériser la texture en couleur.

La deuxième partie présente une architecture de réseaux de neurones fondée sur le principe « diviser pour régner ». Le principe de notre architecture consiste à simplifier la tâche d'un classifieur en divisant le problème à résoudre. L'architecture que nous avons définie permet de simplifier la reconnaissance de données, d'améliorer l'apprentissage, de simplifier la structure des réseaux de neurones utilisés mais principalement d'augmenter le taux de reconnaissance du classifieur sur les données. Nous avons détaillé chacun des points importants de l'architecture MONNA (*Multiple Ordinate Neural Network Architecture*) à savoir sa création, son apprentissage et sa hiérarchisation. Nous avons comparé l'architecture MONNA avec un réseau de neurones multicouche et notre architecture s'est avérée plus performante sur toutes les bases de données que nous avons testées. Afin d'améliorer le taux de reconnaissance de l'architecture MONNA, nous avons également introduit la notion de sélection de caractéristiques qui permet d'une part de simplifier la structure des réseaux de neurones utilisés et d'autre part d'améliorer le taux de reconnaissance de l'architecture.

Ces deux premières parties définissent donc des méthodes qui font partie intégrante d'un système de vision par ordinateur : la segmentation, la caractérisation et la classification. Nous avons défini chacune de ces étapes de façon totalement indépendante afin de permettre une intégration aisée de nouvelles méthodes. Cela permet également de considérer ces méthodes dans un cadre général et elles peuvent être utilisées pour résoudre des problèmes divers de vision par ordinateur.

La troisième partie présente l'élaboration d'un système d'aide au *screening* par le tri informatique cellulaire. Le rôle du système est de compléter l'analyse des pathologistes par des indications nouvelles. Nous avons donc opté pour la construction d'un système semi-automatique servant de contrôle qualité au processus de *screening*. L'élaboration d'un tel système fait appel à des connaissances de traitement d'images, d'intelligence artificielle et de biologie. Le caractère pluridisciplinaire de ce travail est très important et la bonne collaboration entre les experts des différents domaines est déterminante. Le système développé se nomme *Aide à la Recherche en Cytologie par le Tri Informatique Cellulaire* (ARCTIC). Le système ARCTIC fait en quelque sorte la synthèse des méthodes proposées dans les deux premières parties. Nous avons réalisé un système totalement autonome qui permet la reconnaissance des cellules normales et anormales et ceci de façon très fiable. La segmentation des cellules de microscopie est très efficace même pour des cellules difficiles à segmenter (noyaux clairs, cellules en amas). La reconnaissance des objets par l'architecture MONNA s'est avérée très précise puisque 94.5% des cellules anormales et 99% des cellules normales sont reconnues, ce qui est largement supérieur au taux de réussite d'un expert. Au vu des résultats obtenus, la coopération de notre système avec les pathologistes devrait permettre d'éviter la plupart des omissions de cellules anormales.

Cependant certaines améliorations peuvent être envisagées. Tout d'abord nous avons pu constater que la plupart des erreurs de segmentation (qui risquent d'entraîner des erreurs de classification) étaient dues à une mauvaise mise au point lors de l'acquisition de l'image. Une mise au point réalisée de façon automatique semble être la meilleure solution. Cependant cette mise au point ne devra pas consister uniquement à déterminer l'image la plus nette car certaines cellules risquent d'être floues sur l'image même si la majorité des cellules est nette. La mise au point devra consister à construire l'image la plus nette possible à partir d'images obtenues à des distances focales différentes.

Les configurations de cellules en amas ne sont pas extraites de façon optimale par la segmentation. Ce point reste à améliorer. Certains amas présentent des informations intéressantes pour les experts pathologistes, notamment les cellules qui se trouvent en périphérie et qui une fois segmentées et classées peuvent donner des indications sur les amas. Un complément du travail consisterait donc à améliorer ce point et à segmenter de façon plus précise les amas de cellules. Une élaboration de nouvelles caractéristiques portant sur les amas sera également nécessaire car les techniques de caractérisation des amas diffèrent de celles appliquées aux cellules, notamment au niveau de l'échelle d'observation.

Les différentes méthodes de sélection améliorent notablement le taux de reconnaissance de l'architecture MONNA, cependant, il apparaît que l'utilisation d'algorithmes génétiques permet d'obtenir des résultats sensiblement meilleurs. Ces derniers constituent donc une voie prometteuse pour améliorer la sélection de caractéristiques de l'architecture.

Pour terminer, cette thèse fût le fruit de trois années de travail au LUSAC et en collaboration avec le GREYC et le service d'anatomie et cytologie pathologiques de l'hôpital Pasteur de Cherbourg). La collaboration des experts des différents domaines (Informatique, Image et Médical) a été déterminante pour le déroulement de la thèse. Le système développé reste encore à l'état expérimental mais des nouveaux travaux devraient permettre de le compléter.

Résumé

Dans ce travail, nous avons étudié les différentes étapes utiles à l'élaboration d'un système de vision par ordinateur : la segmentation d'images, la caractérisation d'objets et la classification de données. Nous avons développé des techniques originales suffisamment générales permettant de réaliser chacune de ces étapes.

Nous exposerons tout d'abord une méthode de segmentation d'images couleur de microscopie basée sur une ligne de partage des eaux utilisant des informations locales et globales. Afin de pouvoir adapter les points critiques de la méthode, nous avons défini une stratégie d'optimisation qui permet de définir la façon de les aborder et de les optimiser. Nous nous sommes plus particulièrement intéressés au choix de l'espace couleur et à l'obtention de marqueurs. En particulier, nous avons défini une stratégie de choix de l'espace couleur basée sur une mesure du contraste couleur à partir d'informations *a priori*.

Nous détaillons ensuite une architecture de réseaux de neurones. Son principe consiste à simplifier la tâche du classifieur en divisant le problème à résoudre. L'architecture que nous avons définie permet de simplifier la reconnaissance des données, d'améliorer l'apprentissage, de simplifier la structure des réseaux de neurones utilisés mais principalement d'augmenter le taux de reconnaissance du classifieur. Nous avons illustré ces propriétés par des expérimentations sur différentes bases de données.

Enfin, nous présentons l'élaboration d'un système d'aide au *screening* par le tri informatique cellulaire. Ce système réalise la synthèse des méthodes proposées dans les deux premières parties. Il est totalement autonome et permet la reconnaissance des cellules de façon très fiable (94.5% des cellules anormales et 99% des cellules normales).

Mots-clés : Images ; Traitement des ; Techniques numériques ; Vision par ordinateur ; Couleur ; Morphologie mathématique ; Réseaux de neurones (Informatique) ; Classification Automatique ; Cytologie ;

Abstract

In this work, we studied the various stages useful for the development of a computer vision system : the segmentation of images, the characterization of objects and the classification of data. We developed sufficiently general original techniques allowing to carry out each one of these stages.

We will first of all present a method of segmentation of images color of microscopy based on a watershed using of local and global information. In order to be able to adapt the critical points of the method, we defined a strategy of optimization which makes it possible to define the way to approach and to optimize them. We are more particularly interested in the choice of color space and in the obtaining of the markers. In particular, we defined a strategy for the choice of a color space using a color contrast measure based on *a priori* information.

We detail then a neural network architecture. Its principle consists in simplifying the task of the classifier by dividing the problem to be solved. Architecture that we defined makes it possible to simplify the recognition of the data, to improve the training, to simplify the structure of the neural networks used but mainly to increase the recognition rate of the classifier. We illustrated these properties by experiments on various data bases.

Finally, we present the development of a screening assistance system by computer cell sorting. This system carries out the synthesis of the methods suggested in the first two parts. It is completely autonomous and allows the recognition of the cells in a highly reliable way (94.5% of the abnormal cells and 99% of the normal cells).

Keywords : Images ; Processing of ; Numerical Techniques ; Computer Vision ; Color ; Mathematical Morphology ; Neural Networks (Computer Science) ; Automatic classification ; Cytology